

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2005, 18:40:12 ; Search time 161 Seconds
(Without alignments)
248.086 Million cell updates/sec

Title: US-09-763-393-1

Perfect score: 541
Sequence: 1 MSNRVSRSRGRGDGEAPD.....EKTPPKAKTKYKADGQP 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10E_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10F_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US11A_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	541	100.0	102	US-10-205-823-186	Sequence 186, App
2	541	100.0	102	US-10-821-273-2	Sequence 2, Appl
3	528	97.6	99	US-10-131-409-46	Sequence 46, Appl
4	528	97.6	99	US-10-139-854-46	Sequence 46, Appl
5	528	97.6	99	US-10-150-813-46	Sequence 46, Appl
6	528	97.6	99	US-10-150-811-46	Sequence 46, Appl
7	295.5	54.6	109	US-10-131-409-28	Sequence 28, Appl
8	295.5	54.6	109	US-10-131-409-38	Sequence 38, Appl
9	295.5	54.6	109	US-10-139-854-28	Sequence 28, Appl
10	295.5	54.6	109	US-10-139-854-38	Sequence 38, Appl
11	295.5	54.6	109	US-10-150-813-28	Sequence 28, Appl

12	295.5	54.6	109	US-10-150-813-38	Sequence 38, Appl
13	295.5	54.6	109	US-10-150-811-28	Sequence 28, Appl
14	295.5	54.6	109	US-10-150-811-38	Sequence 38, Appl
15	219	40.5	41	US-10-029-386-31312	Sequence 31312, A
16	181.5	33.5	111	US-10-054-683-37	Sequence 37, Appl
17	181.5	33.5	111	US-10-131-409-4	Sequence 4, Appl
18	181.5	33.5	111	US-10-131-409-37	Sequence 37, Appl
19	181.5	33.5	111	US-10-139-854-4	Sequence 4, Appl
20	181.5	33.5	111	US-10-139-854-37	Sequence 37, Appl
21	181.5	33.5	111	US-10-150-813-4	Sequence 4, Appl
22	181.5	33.5	111	US-10-150-813-37	Sequence 37, Appl
23	181.5	33.5	111	US-10-150-811-4	Sequence 4, Appl
24	181.5	33.5	111	US-10-150-811-37	Sequence 37, Appl
25	181.5	33.5	111	US-10-473-603-37	Sequence 37, Appl
26	176	32.5	110	US-10-054-683-21	Sequence 21, Appl
27	176	32.5	110	US-10-188-832-143	Sequence 143, Appl
28	176	32.5	110	US-10-473-603-21	Sequence 21, Appl
29	168.5	31.1	111	US-10-131-409-6	Sequence 6, Appl
30	168.5	31.1	111	US-10-131-409-110	Sequence 110, Appl
31	168.5	31.1	111	US-10-139-854-6	Sequence 6, Appl
32	168.5	31.1	111	US-10-139-854-110	Sequence 110, Appl
33	168.5	31.1	111	US-10-150-813-6	Sequence 6, Appl
34	168.5	31.1	111	US-10-150-813-110	Sequence 110, Appl
35	168.5	31.1	111	US-10-150-811-6	Sequence 6, Appl
36	168.5	31.1	111	US-10-150-811-110	Sequence 110, Appl
37	156.5	28.9	137	US-10-072-012-431	Sequence 431, App
38	156	28.8	111	US-09-974-298-127	Sequence 127, App
39	156	28.8	111	US-10-093-766-36	Sequence 36, App
40	156	28.8	111	US-10-072-012-110	Sequence 110, App
41	156	28.8	111	US-10-072-012-112	Sequence 112, App
42	156	28.8	111	US-10-072-012-432	Sequence 432, App
43	156	28.8	111	US-10-072-012-433	Sequence 433, App
44	156	28.8	111	US-10-473-603-39	Sequence 39, Appl
45	147.5	27.3	185	US-10-072-012-434	Sequence 434, App

ALIGNMENTS

RESULT 1
US-10-205-823-186
Sequence 186, Application US/10205823
Publication No. US20030108963A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Endege, Wilson O.
APPLICANT: Gannavarapu, Manjula
APPLICANT: Gorbacheva, Bella
APPLICANT: Hoersch, Sebastian
APPLICANT: Kamatkar, Shubhangt
APPLICANT: Wonesey, Angela M.
APPLICANT: Glatz, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Duettin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-044
CURRENT APPLICATION NUMBER: US/10/205,823
CURRENT FILING DATE: 2002-07-25
PRIOR APPLICATION NUMBER: 60/307,982
PRIOR FILING DATE: 2001-07-25
PRIOR APPLICATION NUMBER: 60/314,356
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/325,020
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/341,746
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: 60/362,158
PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 186
LENGTH: 102
TYPE: PRT
ORGANISM: Homo sapiens
US-10-823-186

Query Match 100.0%; Score 541; DB 14; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSARVSRSGRGDGEAPDVFAVAPGESQOEPPPTNDIIEPGQEREGTPIERKYE 60
DB 1 MSARVSRSGRGDGEAPDVFAVAPGESQOEPPPTNDIIEPGQEREGTPIERKYE 60

QY 61 GDCQEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102
DB 61 GDCQEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102

RESULT 2

US-10-821-2732
Sequence 2 Application US/10821273
Publication No. US20040248256A1

GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Javali, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Metberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steinger II, Robert J.
APPLICANT: Bowman, Michael R.
APPLICANT: Di Blasio-Smith, Elizabeth
APPLICANT: Widom, Angela
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 00766.000101.
CURRENT FILING DATE: US/10/821,273
CURRENT FILING DATE: 2004-04-09
PRIOR APPLICATION NUMBER: US 09/306,111
PRIOR FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: US 60/084,564
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: US 60/087,645
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: US 60/093,712
PRIOR FILING DATE: 1998-07-22
PRIOR APPLICATION NUMBER: US 60/094,935
PRIOR FILING DATE: 1998-07-31
PRIOR APPLICATION NUMBER: US 60/095,880
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: US 60/096,068
PRIOR FILING DATE: 1998-08-11
NUMBER OF SEQ ID NOS: 180
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 2
LENGTH: 102
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-273-2

Handwritten: 180
#2 p. 64-65

Query Match 100.0%; Score 541; DB 16; Length 102;
Best Local Similarity 100.0%; Pred. No. 2.1e-42;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSARVSRSGRGDGEAPDVFAVAPGESQOEPPPTNDIIEPGQEREGTPIERKYE 60
DB 1 MSARVSRSGRGDGEAPDVFAVAPGESQOEPPPTNDIIEPGQEREGTPIERKYE 60

QY 61 GDCQEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102
DB 61 GDCQEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102

RESULT 3
US-10-131-409-46
Sequence 46 Application US/10131409
Publication No. US20030199465A1

GENERAL INFORMATION:
APPLICANT: Malysankar et al.
TITLE OF INVENTION: No. US20030199465A1 Polypeptides and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-675C1P1CON1
CURRENT FILING DATE: US/10/131,409
CURRENT FILING DATE: 2002-10-24
PRIOR APPLICATION NUMBER: 09/896,954
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 60/182,733
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 60/182,724
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 60/183,896
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: 60/184,497
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/224,157
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: 60/184,482
PRIOR FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 60/184,744
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/197,083
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 60/233,405
PRIOR FILING DATE: 2000-09-18
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 135
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 46
LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
US-10-131-409-46

Query Match 97.6%; Score 528; DB 14; Length 99;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 RYRSRSRGDGEAPDVFAVAPGESQOEPPPTNDIIEPGQEREGTPIERKYE 63
DB 1 RYRSRSRGDGEAPDVFAVAPGESQOEPPPTNDIIEPGQEREGTPIERKYE 60

QY 64 QEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102
DB 61 QEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 99

RESULT 4
US-10-139-854-46
Sequence 46 Application US/10139854
Publication No. US20030202971A1
GENERAL INFORMATION:
APPLICANT: Majumder, Kumud
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-675CON2
CURRENT FILING DATE: US/10/139,854
CURRENT FILING DATE: 2002-12-02
PRIOR APPLICATION NUMBER: 09/783,429
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/182,733
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 60/182,724
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 60/183,896
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: 60/184,497

```

; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; PRIOR FILING DATE: 2000-09-18
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-139-854-46
```

```

Query Match      97.6%; Score 528; DB 15; Length 99;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      4 RVSRSRGRGDDGEADPVVAFVAPGESQOEPPPTNDIDEPQERGTPIIERKYEGDC 63
          |||
DB      1 RVSRSRGRGDDGEADPVVAFVAPGESQOEPPPTNDIDEPQERGTPIIERKYEGDC 60

QY      64 OEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGGQGP 102
          |||
DB      61 OEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGGQGP 99
```

```

RESULT 5
US-10-150-813-46
; Sequence 46, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
; APPLICANT: Malyanckar
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675CIP1CON2
; CURRENT APPLICATION NUMBER: US/10/150,813
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 09/898,954
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/182,733
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-150-813-46
```

Query Match 97.6%; Score 528; DB 15; Length 99;

```

Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 RVSRSRGRGDDGEADPVVAFVAPGESQOEPPPTNDIDEPQERGTPIIERKYEGDC 63
          |||
DB      1 RVSRSRGRGDDGEADPVVAFVAPGESQOEPPPTNDIDEPQERGTPIIERKYEGDC 60

QY      64 OEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGGQGP 102
          |||
DB      61 OEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGGQGP 99
```

```

RESULT 6
US-10-150-811-46
; Sequence 46, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
; APPLICANT: Malyanckar et al.
; TITLE OF INVENTION: No. US20040010120A1 Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/150,811
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 09/970,607
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/182,733
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/182,724
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 60/183,896
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/184,497
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/224,157
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/184,482
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/184,744
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/197,083
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/233,405
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-150-811-46

Query Match      97.6%; Score 528; DB 15; Length 99;
Best Local Similarity 100.0%; Pred. No. 3.3e-41;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 RVSRSRGRGDDGEADPVVAFVAPGESQOEPPPTNDIDEPQERGTPIIERKYEGDC 63
          |||
DB      1 RVSRSRGRGDDGEADPVVAFVAPGESQOEPPPTNDIDEPQERGTPIIERKYEGDC 60

QY      64 OEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGGQGP 102
          |||
DB      61 OEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGGQGP 99

RESULT 7
US-10-131-409-28
; Sequence 28, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
; APPLICANT: Malyanckar et al.
; TITLE OF INVENTION: No. US20030199465A1 Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-675CIP1CON1
; CURRENT APPLICATION NUMBER: US/10/131,409
```


Thu Aug 25 08:49:42 2005

us-09-763-393-1.rapb

Page 7

Oy 57 RKVEGDCQEMDLKPTRSERGDGSDVKEKTPNPYKAATKEA 97
| | | | | | | | | | | | | | | | | | | | | |
D6 1 RKVEGDCQEMDLKPTRSERGDGSDVKEKTPNPYKAATKEA 41

Search completed: August 24, 2005, 18:46:36
Job time : 169 sec

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2005, 18:30:13 ; Search time 58 Seconds

(Without alignments)
900.553 Million cell updates/sec

Title: US-09-763-393-1

Sequence: 1 MSARVSRSGRGDGGAPD.....EKTPPNKTKAKTGADGAP 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : uniprot_03:*
1: uniprot_sprot:*
2: uniprot_tramb1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	541	100.0	102	1 GGCI_HUMAN	060829 homo sapien
2	176	32.5	130	2 Q96GUL	096911 homo sapien
3	168.5	31.1	111	2 Q722X7	072247 homo sapien
4	163.5	30.2	108	1 GGD5_HUMAN	08WMM1 homo sapien
5	156	28.8	111	1 GGD4_HUMAN	08WCP9 homo sapien
6	133	24.6	146	1 GGB1_HUMAN	075459 homo sapien
7	131.5	24.3	111	1 GGD3_HUMAN	096959 homo sapien
8	131	24.2	117	2 Q6NT33	06NT33 homo sapien
9	127	23.5	116	1 GGB8_HUMAN	09UEU5 homo sapien
10	123	22.7	117	1 GGB5_HUMAN	013068 homo sapien
11	123	22.7	118	1 GGB3_HUMAN	013067 homo sapien
12	122	22.6	117	1 GGB2_HUMAN	013070 homo sapien
13	120	22.2	116	1 GGB2_HUMAN	013066 homo sapien
14	120	22.2	116	2 Q6NT46	06NT46 homo sapien
15	120	21.6	138	1 GGB7_HUMAN	013065 homo sapien
16	117	21.4	117	1 GGB7_HUMAN	076087 homo sapien
17	116	21.4	117	1 GGB7_HUMAN	081YCS homo sapien
18	113	20.7	139	2 Q8NA60	08NA60 homo sapien
19	112	20.7	829	2 Q8NA60	095404 homo sapien
20	112	20.7	861	2 Q8NA60	095404 homo sapien
21	112	20.7	861	2 Q8NA60	095404 homo sapien
22	112	20.7	861	2 Q8NA60	095404 homo sapien
23	108.5	20.1	485	2 Q6NXX0	06NXX0 mus musculu
24	108.5	20.1	485	2 Q6NXX0	06NXX0 mus musculu
25	104.5	19.3	792	2 Q9KFO3	09KFO3 bacillus ha
26	102.5	18.9	69	2 Q8WMM0	08WMM0 homo sapien
27	96.5	17.8	229	1 VBA_HPV08	066425 human papil
28	96.5	17.8	325	2 Q44016	044016 dictyostell
29	94.5	17.5	524	1 T2FA_XENLA	062PT3 mus musculu
30	94.5	17.5	524	1 T2FA_XENLA	062PT3 mus musculu
31	94.5	17.5	524	2 Q8JGL7	08JGL7 xenopus lae

32	94	17.4	284	2 Q8HYJ0	08HYJ0 felis silve
33	92.5	17.1	197	2 Q23872	023872 dictyostell
34	92.5	17.1	555	2 Q8NA99	08NA99 homo sapien
35	92.5	17.1	686	2 Q6ZRC5	06ZRC5 homo sapien
36	92.5	17.1	780	2 Q96C66	096C66 homo sapien
37	92.5	17.1	1783	2 Q15038	015038 homo sapien
38	92.5	17.1	1791	2 Q60382	060382 homo sapien
39	92.5	17.1	2286	2 Q9UHA8	09UHA8 homo sapien
40	92	17.0	606	1 CENB_CRIGR	09UHA8 homo sapien
41	92	17.0	702	2 Q7TYX8	07TYX8 cricetulus
42	92	17.0	910	2 Q7TQ48	07TQ48 mus musculu
43	91.5	16.9	727	2 Q95TG7	095TG7 drosophila
44	91.5	16.9	5303	2 Q9V628	09V628 drosophila
45	91	16.8	2480	1 RPL1_HUMAN	081W17 homo sapien

ALIGNMENTS

RESULT 1	GGCI_HUMAN	STANDARD;	PRT;	102 AA.
AC	060829;			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	G antigen family C 1 protein (Prostate-associated gene protein 4)			
DE	(PAGE-4) (PAGE-1) (JM27) (GAGE-9).			
GN	Name=GAGEC1; Synonym=GAGE4;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98393718; PubMed=9724777; DOI=10.1073/pnas.95.18.10757;			
RA	Brinkmann U., Vasmatazis G., Lee B., Yernool N., Essand M.,			
RA	Pastan I.,			
RT	"PAGE-1, an X chromosome-linked GAGE-like gene that is expressed in			
RL	normal and neoplastic prostate, testis, and uterus."			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:10757-10762(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Strom T.M., Nyakatura G., Heilbrand H., Dreescher B., Rosenthal A.,			
RA	Weinl A.,			
RN	Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L.H., Shennan C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buettow K.H., Scheffer C.F., Heist N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist N.K.,			
RA	Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Panje C.,			
RA	Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosch S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richard S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield V.S.N., Krzywicki M.I., Skalska U., Smallus D.E.,			
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.,			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
CC	-1- FUNCTION: Unknown.			
CC	-1- TISSUE SPECIFICITY: Preferentially expressed in normal male and			
CC	female reproductive tissues, prostate, testis, fallopian tube,			

Sgt 1/18

```
CC      uterus, and placenta, as well as in prostate cancer, testicular
CC      cancer, and uterine cancer.
CC      -1- SIMILARITY: Belongs to the GAGE family.
CC      -----
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF275258; AAF88037.1; -.
DR      EMBL; AJ005894; CA06751.1; -.
DR      EMBL; AF238380; -. NOT ANNOTATED_CDS.
DR      EMBL; BC010897; AAH0897.1; -.
DR      H-InvDB; HIX0016804; -.
DR      Genew; HGNC:4108; GAGEC1.
DR      InterPro; IPR008625; GAGE.
DR      Pfam; PF05831; GAGE; 1.
KW      Multigene family.
SQ      SEQUENCE 102 AA; 11153 MW; CESD07AFBF73301B CRC64;

Query Match          100.0%; Score 541; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.6e-36;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSARVRSRGRGDGGADPVVAFAVAPGSGQGEPPPTNDQIEPQERGTPIERKYE 60
DB      1 MSARVRSRGRGDGGADPVVAFAVAPGSGQGEPPPTNDQIEPQERGTPIERKYE 60
QY      61 GDCQEMDLKTRSERGSDGVKKTPTPNPKATKXAGGQP 102
DB      61 GDCQEMDLKTRSERGSDGVKKTPTPNPKATKXAGGQP 102

RESULT 2
Q96GCU1 PRELIMINARY; PRT; 130 AA.
AC      Q96GCU1; Q8WWL9;
DT      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT      25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE      PAGE-5 protein.
GN      Name=PAGE-5;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Skin;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ueda T.B., Toshiyuki S., Carinici P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA      Jones S.J., Marra M.A.,
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
```

```
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Skin;
RA      Struhsberg R.;
RL      Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Zeman A.J.W., van Kraats A.A., Weidle U.H., Ruiter D.J.,
RA      Van Muljen G.N.P.;
RL      Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC009230; AAH09230.2; -.
DR      EMBL; AJ344352; CAC87642.1; -.
DR      InterPro; IPR008625; GAGE.
DR      Pfam; PF05831; GAGE; 1.
SQ      SEQUENCE 130 AA; 14046 MW; 17560A12BE103424 CRC64;

Query Match          32.5%; Score 176; DB 2; Length 130;
Best Local Similarity 43.1%; Pred. No. 6.3e-07;
Matches 47; Conservative 11; Mismatches 43; Indels 8; Gaps 3;

QY      1 MSARVRSRGRGDGQA-----PDVVAFAVAPGSGQGEPPPTNDQIEPQGE--RGTP 52
DB      21 MSEHYTRSGSERGNDGSSQPVGPVIVQPTTEKRGQEEPPPTNDQIAPSGEIKNEGAP 80
QY      53 PIEERKVGDDQEMDLKTRSERGSDGVKKTPTPNPKATKXAGGQP 101
DB      81 AVQGTVAEPQDELALTKIEDAPGDPVREGTPTFPPTVVLAGEGQ 129

RESULT 3
Q722X7 PRELIMINARY; PRT; 111 AA.
AC      Q722X7;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      Similar to PAGE-5 protein.
GN      Name=MG62094 (Human).
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxId=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Poolied;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Ueda T.B., Toshiyuki S., Carinici P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA      Boesak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Krzywinski M.I., Skalska U., Smalls D.E., Scherch A., Schein J.E.,
RA      Jones S.J., Marra M.A.,
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Poolied;
RA      Struhsberg R.;
RL      Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC054022; AAH54022.1; -.
```

DR InterPro; IPR008625; GAGE.
 DR EMBL; AF05831; GAGE; 1.
 SO SEQUENCE 111 AA; 12090 MW; C0692E3DP2C905F8 CRC64;

Query Match 31.1%; Score 168.5; DB 2; Length 111;
 Best Local Similarity 42.5%; Pred. No. 2.1e-06;
 Matches 45; Conservative 10; Mismatches 44; Indels 7; Gaps 3;

QY 4 RVRSRGRGDCGEADPVAVFVAFGE-----SQGEPPPTNDIIEP-GQ-EREGTPIIEE 56
 DB 6 RAASQSSERNDDSSQPVGSVIVQEPTEKRCQEEPPPTNDQGIASGEIENQAVPAFQG 65
 Db 57 RAYEGCGQEMDLEKTRSERGSDGVKKEPTPNPKHAKTKXAGGCGP 102
 66 PDMEARQOGLALAKIEDPFGDPVREGIMPTDILTKVLEAGDAQP 111

RESULT 4
 GCD5 HUMAN STANDARD; PRT; 108 AA.
 AC 08KMTL;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE G antigen family D 5 protein (XAGE-5 protein).
 GN Name=GAGED5; Synonyms=XAGE5;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22021429; PubMed=11992404; DOI=10.1002/jgc.10371;
 RA Zenden A.J.W., Van Kraats A.A., Weidie U.H., Ruiters D.J.,
 RA Van Muljen G.N.P.;
 RT "The XAGE family of cancer/testis-associated genes: alignment and
 RT expression profile in normal tissues, melanoma lesions and Ewing's
 RT sarcoma.";
 RT Int. J. Cancer 99:361-369(2002).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko U., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RL [3]
 RN -1- FUNCTION: Unknown.
 -1- SIMILARITY: Belongs to the GAGE family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL; AJ318894; CAC83015.1; -;
 DR EMBL; BC069129; AAH69129.1; -;
 DR Genew; HGNC:30930; GAGED5.
 DR InterPro; IPR008625; GAGE.
 DR EMBL; AF05831; GAGE; 1.
 DR Pfam; PF05831; GAGE; 1.
 KM Multigene family.
 SO SEQUENCE 108 AA; 12077 MW; E3CD91E5C9241628 CRC64;

Query Match 30.2%; Score 163.5; DB 1; Length 108;
 Best Local Similarity 45.5%; Pred. No. 5.3e-06;
 Matches 35; Conservative 10; Mismatches 29; Indels 3; Gaps 1;

QY 29 ESQGEPPPTNDIIEPGGSE---GPPPEERKRVGDCQEMDLEKTRSERGCGSDVKEKT 85
 DB 30 EPQGEPPPTESQDHTFGKREDDQGAETQVPLNLEADQLSOSKTGDECGSDPVQGKI 89
 Db 86 PPNPKHAKTKXAGDQGP 102
 90 LPKSBOFKMPEGGEGKPF 106

RESULT 5
 GCD4 HUMAN STANDARD; PRT; 111 AA.
 ID GCD4 HUMAN
 AC 08MTD9; 08MY59;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE G antigen family D 4 protein (Placenta-specific gene 6 protein) (XAGE-
 DE 3 protein) (ps9012).
 GN Name=GAGED4; Synonyms=PLAC6, XAGE3;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_Taxid=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22021429; PubMed=11992404; DOI=10.1002/jgc.10371;
 RA Zenden A.J.W., Van Kraats A.A., Weidie U.H., Ruiters D.J.,
 RA Van Muljen G.N.P.;
 RT "The XAGE family of cancer/testis-associated genes: alignment and
 RT expression profile in normal tissues, melanoma lesions and Ewing's
 RT sarcoma.";
 RT Int. J. Cancer 99:361-369(2002).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP Zhou X.M., Zhang P.F., Jiang H.Q., Huang Y., Qin W.X., Zhao X.T.,
 RA Wan D.F., Gu J.R.;
 RT "Novel human cDNA clones with function of inhibiting cancer cell
 RT growth.";
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RL [3]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Placenta;
 RX MEDLINE=22288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko U., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Unknown.
 CC -1- SIMILARITY: Belongs to the GAGE family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AJ318881; CAC8308.1; -;
 CC DR EMBL; AJ318893; CAC83014.1; -;
 CC DR EMBL; AF318372; AAL55879.1; -;
 CC DR EMBL; BC062680; AAB62680.1; -;
 CC DR Genew; HGNC:14618; GAGED4.
 CC DR InterPro; IPR008625; GAGE.
 CC DR Pfam; PF05831; GAGE; 1.
 CC KM Multigene family.
 CC FT CONFLICT 11 11 P -> H (in Ref. 2).
 SQ SEQUENCE 111 AA; 12302 MW; F4A3261B6F7ACF5B CRC64;
 Query Match 28.8%; Score 156; DB 1; Length 111;
 Best Local Similarity 41.1%; Pred. No. 2,2e-05;
 Matches 37; Conservative 13; Mismatches 34; Indels 6; Gaps 3;
 QY 19 PDVVA-FVAPG--ESQOEPPPTNDQIEPQERE--GTPPIERKVEGDCQMDLEKTR 72
 DB 20 PELIGMLEPGEDEEPOEPEPTESRPAQGEREDQGAETQVPLTADLDELQSQTG 79
 QY 73 SERGSDGVKPKTPPKPKAKTKTEADGAP 102
 DB 80 GEGCNGPDQCKLPLKSEQPKPEGGDRDP 109

RESULT 6
 GGB1_HUMAN STANDARD; PRT; 146 AA.
 AC 075459; Q9BS87;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE G antigen family B.1 protein (Prostate-associated gene protein 1)
 DE (PAGE-1) (GAGE-9) (A45).
 GN Name=GAGEB1; Synonyms=GAGE9, PAGE1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98316329; PubMed=9651357; DOI=10.1074/jbc.273.28.17618;
 RA Chen M.B., Lin S.-H., Chung L.W.K., Sikes R.A.;
 RT "Isolation and characterization of PAGE-1 and GAGE-7: new genes
 RT expressed in the LNCaP prostate cancer progression model that share
 RT homology with melanoma-associated antigens.";
 RL J. Biol. Chem. 273:17618-17625(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins S.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stalcup M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.D., Ustin T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullishy S.J.,
 RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultine S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Unknown.
 CC -1- TISSUE SPECIFICITY: Isolated from prostate cancer cell lines;
 CC expression associated with progression to androgen insensitive
 CC phenotype. Expressed in normal testis and at lower level in normal
 CC placenta.
 CC -----
 CC -1- SIMILARITY: Belongs to the GAGE family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF058989; AAC25990.1; -;
 CC DR EMBL; BC004861; AAB04861.1; -;
 CC DR Genew; HGNC:4107; GAGEB1.
 CC DR H-InvDB; HIX0016803; -;
 CC DR MIM; 300288; -;
 CC DR GO; GO:0006968; P:cellular defense response; TAS.
 CC DR InterPro; IPR008625; GAGE.
 CC DR Pfam; PF05831; GAGE; 1.
 CC KM Multigene family.
 CC FT CONFLICT 75 75 P -> L (in Ref. 2).
 SQ SEQUENCE 146 AA; 16134 MW; 91ABA1E3B49BDC42 CRC64;
 Query Match 24.6%; Score 133; DB 1; Length 146;
 Best Local Similarity 33.6%; Pred. No. 0.0021;
 Matches 38; Conservative 13; Mismatches 40; Indels 22; Gaps 3;
 QY 7 SRSRGDQGAQDVVAFVAPGESQOEPPPTNDQI-----ERG-----OER 48
 DB 35 TQSGSTPAEBREDEGASAGG---QEPAESQELVPTKGCEDGDPPTKVCRLRBE 90
 QY 49 ECTPIERKVEGDCQEMDLKTRSERGSDGVKPKPKAKTKTEADGAGGQ 101
 DB 91 QMKLPABEPPEPADSQEQVHPKTCGERGDPVQELGPNPEEVKTPPEDEGQ 143

RESULT 7
 GGD3_HUMAN STANDARD; PRT; 111 AA.
 AC Q96GT9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE G antigen family D.3 protein (XAGE-2 protein).
 DE Name=GAGED3; Synonyms=XAGE2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99211468; PubMed=10197611;
 RA Brinkmann U., Vaematzis G., Lee B., Pastan I.;
 RT "Novel genes in the PAGE and GAGE family of tumor antigens found by
 RT homology walking in the dbEST database.";
 RL Cancer Res. 59:1445-1448(1999).
 RL [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22021429; PubMed=11992404; DOI=10.1002/jc.10371;

RA Zendenman A.J.W., Van Kreats A.A., Weidle U.H., Ruiters D.J.,
 RA Van Wulfen G.N.P.;
 RT "The XAGE family of cancer/testis-associated genes: alignment and
 RT expression profile in normal tissues, melanoma lesions and Ewing's
 RT sarcoma.";
 RL Int. J. Cancer 99:361-369(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -|- FUNCTION: Unknown.
 CC -|- SIMILARITY: Belongs to the GAGE family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@1sb-sib.ch).
 CC -----
 CC EMBL/ AJ318891; CAC88125.1; -;
 CC EMBL/ AJ318892; CAC88125.1; JOINED.
 CC EMBL/ AJ318880; CAC83007.1; -;
 CC EMBL/ BC009232; AAH09232.1; -;
 CC Genew; HGNC:4112; GAGED3.
 CC MIM; 300416; -;
 CC InterPro; IPR008625; GAGE.
 CC Pfam; PF05831; GAGE; 1.
 CC MultiGene family.
 CC KW SEQUENCE 111 AA; 12354 MW; 38C93332C5BA0E14 CRC64;
 SQ

Query Match 24.3%; Score 131.5; DB 1; Length 111;
 Best Local Similarity 37.0%; Pred. No. 0.0021;
 Matches 40; Conservative 16; Mismatches 45; Indels 7; Gaps 3;

QY 1 MSARVRS--RSRRGSGQEAADVAVAP--GSGQSEPTNDIIPGQERE--GTPP 53
 DB 1 MSWRGSRSTYPRPRRSIQPELIGAMLEPDEBPKEKPKPTKSRNPPTDQKREDDGAAE 60
 QY 54 IEBRKATGDCQEMDLEKTRSERGDSGVYKTPRPNKATKRAAGGQ 101
 DB 61 IQVPDLADLQELCQRTKTDGCGEGTDVKGKILPKAHFMPKAPAGGK 108

RESULT 8
 Q6NT33 PRELIMINARY; PRT; 117 AA.
 AC Q6NT33;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE GAGE4 protein.

GN Name=GAGE4;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PCR rescued clones;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.B.,
 RA Jones S.J., Matra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC [2]
 CC SEQUENCE FROM N.A.
 RC TISSUE=PCR rescued clones;
 RA Strausberg R.;
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC069470; AAH69470.1; -;
 DR InterPro; IPR008625; GAGE.
 DR Pfam; PF05831; GAGE; 1.
 SQ SEQUENCE 117 AA; 12885 MW; 365BD71B24DCTAD CRC64;
 Query Match 24.2%; Score 131; DB 2; Length 117;
 Best Local Similarity 40.5%; Pred. No. 0.0024;
 Matches 32; Conservative 7; Mismatches 38; Indels 2; Gaps 1;

QY 25 VAPGSGQSEPTNDIIPGQ--EREGTPPIERKRVGDCQMDLEKTRSERGDSGVK 82
 DB 36 VEPATPEBGEPAVYORDDPAAGBEGDEGASAGGPKPEADSOEGHPGTGCECEDEPDG 95
 QY 83 EKTTPNPKHAKTKREAGDGQ 101
 DB 96 EMDPPEBVKTPBEGGQ 114

RESULT 9
 GGE8 HUMAN
 ID GGE8 HUMAN STANDARD; PRT; 116 AA.
 AC Q9UE05;
 DT 28-FEB-2003 (rel. 41, Created)
 DT 28-FEB-2003 (rel. 41, Last sequence update)
 DT 05-JUL-2004 (rel. 44, Last annotation update)
 DE GAGE-8 protein (G antigen 8).
 GN Name=GAGE8;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=9323388; PubMed=10397259;
 RA De Backer O., Arden K.C., Boretti M., Vantomme V., De Smet C.,
 RA CzeKay S., Vliers C.S., De Plessen E., Brasseur F., Chomez P.,
 RA Van den Eynde B., Boon T., Van der Bruggen P.;
 RT "Characterization of the GAGE genes that are expressed in various
 RT human cancers and in normal testis.";

```

RL. Cancer Res. 59:3157-3165(1999).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Pange C.,
RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heltan E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Not expressed in normal tissues, except in
CC testis, but expressed by a large proportion of tumors of various
CC histological origins.
CC -1- SIMILARITY: Belongs to the GAGE family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF055473; AAC33676.1; -.
DR EMBL: BC018052; AAH18052.1; -.
DR Genew; HGNC:4106; GAGE8.
DR GO; GO:0006968; P:cellular defense response; TAS.
DR InterPro; IPR008625; GAGE.
DR Pfam; PF05831; GAGE; 1.
KW Multigene family.
SQ SEQUENCE 116 AA; 12764 MW; DD305239566F19A CRC64;

Query Match 23.5%; Score 127; DB 1; Length 116;
Best Local Similarity 35.4%; Pred. No. 0.0051;
Matches 40; Conservative 10; Mismatches 51; Indels 12; Gaps 3;

QY 1 MSARVRSRRGDCQ-EPADVVA-----FVAFGSQGEPPPTNDQIEPQ--PR 48
DB 1 MSWRGSTRPRPRRVRVEPEMIGMPRPQFSDVEVPATPEEPTQGRDPAAGEGD 60
QY 49 ECTPPIERKVGDCQEMDLEKTRSERGDSVKEKTPPNPKHAKTKKAGDCQ 101
DB 61 EGASAGQKPRPADSGQGHPTQGCCEBPDGQENDPPNPBEVKTPEEGEQ 113

RESULT 10
GGG4 HUMAN STANDARD; PRT; 117 AA.
AC 013068;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 05-JUL-2004 (rel. 44, Last annotation update)
DE GAGE-4 protein (G antigen 4).
GN Name=GAGE4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma."
RL J. Exp. Med. 182:689-698(1995).
CC -1- TISSUE SPECIFICITY: Expressed in a variety of tumor tissues but
CC not in normal tissues, except testis.
CC -1- SIMILARITY: Belongs to the GAGE family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U19145; AAA82747.1; -.
DR Genew; HGNC:4101; GAGE4.
DR MIM; 604246; -.
DR InterPro; IPR008625; GAGE.
DR Pfam; PF05831; GAGE; 1.
KW Multigene family.
SQ SEQUENCE 117 AA; 12956 MW; 365ED71B2F9DC7AD CRC64;

Query Match 22.7%; Score 123; DB 1; Length 117;
Best Local Similarity 39.2%; Pred. No. 0.011;
Matches 31; Conservative 7; Mismatches 39; Indels 2; Gaps 1;

QY 25 VAPGSGQGEPPPTNDQIEPQ--ERETPIERKVGDCQEMDLEKTRSRGDSVYK 82
DB 36 VEPATPEEGEPTQGRDPAAGEGDEKASAGQKPRPADSGQGHPTQGCCEBPDGQ 95
QY 83 EKTTPNPKHAKTKKAGDCQ 101
DB 96 EMDPPNPBEVKTPEEGEQ 114

RESULT 11
GGG5 HUMAN STANDARD; PRT; 117 AA.
AC 013069;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 05-JUL-2004 (rel. 44, Last annotation update)
DE GAGE-5 protein (G antigen 5).
GN Name=GAGE5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378788; PubMed=7544395;
RA van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
RT cytolytic T lymphocytes on a human melanoma."
RL J. Exp. Med. 182:689-698(1995).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

```

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rahn S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Boeck S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- TISSUE SPECIFICITY: Expressed in a variety of tumor tissues but
CC not in normal tissues, except testis.
CC -1- SIMILARITY: Belongs to the GAGE family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: U19146; AAA82748.1; -.
DR EMBL: BC024914; AAA24914.1; -.
DR Genew: HGNC:4102; GAGE5.
DR H-InvDB: HIX002115; -.
DR MIM: 604247; -.
DR InterPro: IPR008625; GAGE.
DR Pfam: PF05831; GAGE; 1.
DR Multigene family.
KW SEQUENCE 117 AA; 12924 MW; 234A65E2ECDD06 CRC64;

Query Match 22.7%; Score 123; DB 1; Length 117;
Best Local Similarity 39.2%; Pred. No. 0.011;
Matches 31; Conservative 7; Mismatches 39; Indels 2; Gaps 1;

OY 25 VAPESQOEPPPTNODLEPQ--EREPTPIERKRYGDCQEMDLEKTSERGDGSDVK 82
DB 36 VEPPTPEGEPPATRODPAQAQESGDEGASAGQPKPEADLSQEOGHPOTGCECEDGPDGQ 95
OY 83 EKTTPNPKHAATKKEAGDQG 101
DB 96 EMDPPNPPEVVTPEGEKQ 114

RESULT 12
GAGE3 HUMAN STANDARD; PRT; 118 AA.
AC Q13067;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE GAGE-3 protein (G antigen 3).
GN Name=GAGE3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma; PubMed=7544395;
RX MEDLINE=95378788;
RA van den Eynde B., Peeters O., de Becker O., Gaugler B., Lucas S.,
RA Boon T.;
RT "A new family of genes coding for an antigen recognized by autologous
cytolytic T lymphocytes on a human melanoma."
J. Exp. Med. 182:689-698(1995).
[2]

```

```

RP PHOSPHORYLATION SITE TYR-10.
RA PubMed=12522270; DOI=10.1073/pnas.2436191100;
RX Salomon A.R., Pizarro S.B., Brill L.M., Brinker A., Phung Q.T.,
RT Erlanson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;
RA "Profiling of tyrosine phosphorylation pathways in human cells using
RT mass spectrometry.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:443-448(2003).
CC -1- TISSUE SPECIFICITY: Expressed in a variety of tumor tissues but
CC not in normal tissues, except testis.
CC -1- SIMILARITY: Belongs to the GAGE family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: U19144; AAA82746.1; -.
CC DR Gene; HGNC:4100; GAGE3.
CC MIM: 604245; -.
CC DR InterPro: IPR006625; GAGE.
CC DR Pfam: PF05831; GAGE, 1.
CC DR Multigene Family: Phosphorylation.
CC KW MOD RES 10 10 Phosphotyrosine.
CC FT SEQUENCE 118 AA; 12937 MW; D97EBB319E725103 CRC64;
SQ
Query Match 22.7%; Score 123; DB 1; Length 118;
Best Local Similarity 39.2%; Pred. No. 0.011;
Matches 31; Conservative 7; Mismatches 39; Indels 2; Gaps 1;
OY 25 VAPGSGOQEPPTNODIEPQ--ERGSTPIERKAYGDCQMDLEKTRSERGDSGVK 82
DB 37 VEPATPREGPATYRQDPAAQSEDEGASAGQSPKPEADSQEGHPTGCECEDPQ 96
OY 83 EKTPEPFGHAKTKKAGDQ 101
DB 97 EMDPPNPEVKTPEGEKQ 115
-----
RESULT 13
GAGE6_HUMAN STANDARD; PRT; 117 AA.
ID ID GAGE6_HUMAN STANDARD; PRT; 117 AA.
AC Q13070;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE GAGE-6 protein (G antigen 6).
GN Name=GAGE6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95378786; PubMed=7544395;
van den Eynde B., Peeters O., de Backer O., Gaugler B., Lucas S.,
Boon T.;
RA "A new family of genes coding for an antigen recognized by autologous
RT cyclolytic T lymphocytes on a human melanoma.";
RL J. Exp. Med. 183:689-698(1995).
CC -1- TISSUE SPECIFICITY: Expressed in a variety of tumor tissues but
CC not in normal tissues, except testis.
CC -1- SIMILARITY: Belongs to the GAGE family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: U19144; AAA82746.1; -.
CC DR Gene; HGNC:4100; GAGE3.
CC MIM: 604245; -.
CC DR InterPro: IPR006625; GAGE.
CC DR Pfam: PF05831; GAGE, 1.
CC DR Multigene Family: Phosphorylation.
CC KW MOD RES 10 10 Phosphotyrosine.
CC FT SEQUENCE 118 AA; 12937 MW; D97EBB319E725103 CRC64;
SQ

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2005, 18:24:55 ; Search time 71 Seconds

(Without alignments)
555.628 Million cell updates/sec

Title: US-09-763-393-1

Perfect score: 541
Sequence: 1 MSARVRSRRGRGDCGEAPD.....EKTTPNPKAKTKKAGDGP 102

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: geneseqp19808:*
- 2: geneseqp19908:*
- 3: geneseqp20008:*
- 4: geneseqp20018:*
- 5: geneseqp20028:*
- 6: geneseqp20038:*
- 7: geneseqp20038:*
- 8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	541	100.0	102	3	AAV83158	Aay83158 PAGE-4 po
2	541	100.0	102	3	AAV52998	Aay52998 Human sec
3	541	100.0	102	5	AAEL14575	Aael14575 Human dif
4	541	100.0	102	7	ADB75362	Adb75362 Prostate
5	541	100.0	102	8	ADL90096	Adl90096 Human pro
6	528	97.6	99	7	ADG42685	Adg42685 Human PAG
7	528	97.6	99	7	ADJ55754	Adj55754 Peptide h
8	528	97.6	99	8	ADM76592	Adm76592 Human NOV
9	419	77.4	89	2	AAV12447	Aay12447 Human 5'
10	391.5	72.4	87	3	AAV83167	Aay83167 PAGE1 pol
11	377.5	69.8	118	5	ADJ33678	Adj33678 Protein o
12	295.5	54.6	109	3	AAV83168	Aay83168 PAGE2 pol
13	295.5	54.6	109	7	ADG42677	Adg42677 Human PAG
14	295.5	54.6	109	7	ADG42667	Adg42667 Human PAG
15	295.5	54.6	109	7	ADJ55845	Adj55845 Peptide h
16	295.5	54.6	109	7	ADJ55746	Adj55746 Peptide h
17	295.5	54.6	109	8	ADM76574	Adm76574 Human NOV
18	295.5	54.6	109	8	ADM76584	Adm76584 Human NOV
19	219	40.5	41	8	ABO57678	Abos7678 Human gen
20	181.5	33.5	111	4	AAE08581	Aae08581 Human NOV
21	181.5	33.5	111	6	ABJ19255	Abj19255 Human can
22	181.5	33.5	111	7	ADG42643	Adg42643 Human pro
23	181.5	33.5	111	7	ADG42676	Adg42676 Human pro
24	181.5	33.5	111	7	ADJ55745	Adj55745 Peptide f
25	181.5	33.5	111	7	ADJ55716	Adj55716 Human NOV

26	181.5	33.5	111	8	ADH71590	Adh71590 Human pro
27	181.5	33.5	111	8	ADH71588	Adh71588 Human pro
28	181.5	33.5	111	8	ADM78795	Adm78795 Human NOV
29	181.5	33.5	111	8	ADM76550	Adm76550 Human NOV
30	181.5	33.5	111	8	ADM76583	Adm76583 Human NOV
31	176.5	32.6	111	8	ADH71592	Adh71592 Human pro
32	176.5	32.6	111	8	ADH71584	Adh71584 Human pro
33	176	32.5	110	6	ABR48212	AbR48212 Human bla
34	176	32.5	110	6	ABUS6511	Abus6511 Lung canc
35	176	32.5	110	6	ABJ19247	Abj19247 Human can
36	168.5	31.1	111	4	AAE08582	Aae08582 Human NOV
37	168.5	31.1	111	4	AAE08593	Aae08593 Protein e
38	168.5	31.1	111	7	ADG42645	Adg42645 Human pro
39	168.5	31.1	111	7	ADG42749	Adg42749 Huma Full
40	168.5	31.1	111	7	ADJ55718	Adj55718 Human NOV
41	168.5	31.1	111	8	ADH71580	Adh71580 Human pro
42	168.5	31.1	111	8	ADH71582	Adh71582 Human pro
43	168.5	31.1	111	8	ADH71586	Adh71586 Human pro
44	168.5	31.1	111	8	ADM78823	Adm78823 Human NOV
45	168.5	31.1	111	8	ADM76552	Adm76552 Human NOV

ALIGNMENTS

RESULT 1	AAV83158	standard; protein; 102 AA.
ID	AAV83158	
XX	AAV83158;	
AC		
XX		
DT	24-UTL-2000	(first entry)
XX		
DE	PAGE-4 polypeptide.	
XX		
KW	PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;	
KW	uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;	
KW	cytotoxic T lymphocyte; immune response; antibody; drug delivery;	
KW	immunconjugate.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200012706-A1.	
XX		
PD	09-MAR-2000.	
XX		
PF	31-AUG-1999;	99WO-US020046.
XX		
PR	01-SEP-1998;	98US-0098993P.
XX		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PA	Pastan I, Brinkmann U, Vasmatazis G, Lee B;	
XX	WPI, 2000-237869/20.	
XX	N-PSDB; AA293524.	
PT		
PT	Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T	
PT	lymphocyte response and for raising antibodies which can be used to	
PT	detect the presence of PAGE-4 in cell samples or body tissues.	
XX		
PS	Disclosure; Fig 1, 63pp; English.	
XX		
CC	PAGE-4 is a gene preferentially expressed in normal male and female	
CC	reproductive tissues e.g. prostate, testis, fallopian tube, uterus and	
CC	placenta, as well as in prostate cancer, testicular cancer and uterine	
CC	cancer. This expression pattern makes it a target for diagnosis and for	
CC	vaccine based therapy of such neoplasms. An isolated PAGE-4 peptide which	
CC	induces a cytotoxic T lymphocyte response when bound to a major	
CC	histocompatibility complex (MHC) class I molecule or the isolated PAGE-4	
CC	protein can be used in immunogenic compositions to raise a cytotoxic T	
CC	lymphocyte response against cells expressing PAGE-4 including cancer	
CC	cells of the prostate, uterus and testis. The nucleic acids encoding PAGE	

CC -4 or PAGE-4 peptide fragments can also be used in these compositions.
CC Antibodies against PAGE-4 and its peptide fragments can be used in
CC detecting the presence of PAGE-4 and PAGE-4 coding sequences in cell
CC samples or body tissues. The presence of PAGE-4 in tissues which are not
CC related to reproduction can be indicative of the spread of cancerous
CC reproductive tissue. PAGE-4 can also be used to raise antibodies which
CC are then used as the targeting group of immunocongulates comprising
CC toxins used in therapeutic applications. This has applications for drug
CC delivery systems

XX SQ Sequence 102 AA;

Query Match 100.0%; Score 541; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.7e-49;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSARVRSRGRGQGEAPDVVAFVAPGSGQEEPTNODIEPQEREGTPPIERKYE 60
DB 1 MSARVRSRGRGQGEAPDVVAFVAPGSGQEEPTNODIEPQEREGTPPIERKYE 60
QY 61 GDCQEMDEKTRSERGSDVKEKTPPNPKAKTKKAGDGP 102
DB 61 GDCQEMDEKTRSERGSDVKEKTPPNPKAKTKKAGDGP 102

RESULT 2

ID AAY52998 standard; protein; 102 AA.

XX AAY52998;

DT 29-FEB-2000 (first entry)

DE Human secreted protein clone hm365_53 protein sequence SEQ ID NO:2.

XX Human; secreted protein; nutritional; cytokine; cell proliferation;
XX differentiation; immune stimulating; vaccine; suppression;
XX haematopoiesis regulation; tissue growth; activin; inhibin; chemotactic;
XX chemokinesis; haemostatic; thrombolytic; receptor; ligand;
XX anti-inflammatory; cadherin; tumour invasion suppressor;
XX tumour inhibition; gene therapy.

XX Homo sapiens;

XX MO9957132-A1;

PD 11-NOV-1999.

PF 07-MAY-1999; 99WO-US009970.

XX 07-MAY-1998; 98US-0084564P.

PR 02-JUL-1998; 98US-0087645P.

PR 22-JUL-1998; 98US-0093712P.

PR 31-JUL-1998; 98US-0094935P.

PR 10-AUG-1998; 98US-0095880P.

PR 11-AUG-1998; 98US-0096068P.

PR 06-MAY-1999; 99US-00306111.

XX (GENE-) GENETICS INST INC.

XX Jacobo K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Merberg D, Treacy M, Agostino MJ, Steininger RJ, Bowman MR;

XX Dibiase-Smith E, Widom A;

XX WPI; 2000-052937/04.

XX N-PSDB; AA233316.

XX New polynucleotides encoding secreted human proteins, derived from adult

XX placenta, adult retina, fetal brain, fetal.

CC isolated from adult placenta, adult retina, foetal brain, foetal kidney,
CC adult blood, adult brain, adult thyroid, adult bladder, adult neural
CC tissue, adult testes, and adult lymph node CDNA libraries. The human
CC secreted proteins, and the polynucleotides encoding them, are predicted
CC to have biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals. Suggested activities include nutritional activity, cytokine and
CC cell proliferation/differentiation activity, immune stimulating (e.g. as
CC vaccine) or suppressing activity, haematopoiesis regulating activity,
CC tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumour
CC invasion suppressor activity, and tumour inhibition activity. The
CC polynucleotides are also stated to be useful for gene therapy. AA233316
CC to AA233373 encode human secreted proteins, and AAY52998 to AAY53060
CC represent human secreted proteins, given in the present invention

XX SQ Sequence 102 AA;

Query Match 100.0%; Score 541; DB 3; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.7e-49;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSARVRSRGRGQGEAPDVVAFVAPGSGQEEPTNODIEPQEREGTPPIERKYE 60
DB 1 MSARVRSRGRGQGEAPDVVAFVAPGSGQEEPTNODIEPQEREGTPPIERKYE 60
QY 61 GDCQEMDEKTRSERGSDVKEKTPPNPKAKTKKAGDGP 102
DB 61 GDCQEMDEKTRSERGSDVKEKTPPNPKAKTKKAGDGP 102

RESULT 3

ID AAB14575 standard; protein; 102 AA.

XX AAB14575;

DT 21-MAY-2002 (first entry)

DE Human differentially expressed BPH protein JT463726.

XX JT463726; prostate; benign prostatic hyperplasia; BPH; human; GAGB-like;

XX PAGE-like.

XX Homo sapiens.

XX WO200210338-A2.

PD 07-FEB-2002.

PF 31-JUL-2001; 2001WO-US023861.

PR 31-JUL-2000; 2000US-0222039P.

PR (GENE-) GENE LOGIC INC.

XX Munger WE, Kulkarni P, Getzenberg RH;

XX WPI; 2002-206184/26.

XX N-PSDB; AAD24228.

XX Method for diagnosing prostatic diseases in a patient, e.g. benign

XX prostatic hyperplasia, or screening compounds for treating these

XX diseases, comprises assaying for the expression of JT463726 protein or

XX nucleic acid in a sample.

XX Claim 17; Page 48; 48pp; English.

XX The invention relates to a method of diagnosing prostatic disease in a

XX patient, which comprises assaying for the expression of JT463726 protein

XX or nucleic acid in a sample. The differential expression of the protein

XX or nucleic acid compared to normal prostate tissue is an indication of


```

Db      1 MSARVSRSRGRGDGEAPDVFAVAPGESQOEPPDNDIPRQEREGTPPIERKVE 60
QY      61 GDCQEMDLEKTRSERGSDVYKKTPTPNPKAKTKRAGDGQP 102
        61 GDCQEMDLEKTRSERGSDVYKKTPTPNPKAKTKRAGDGQP 102
Db
RESULT 6
ADG42685
ID      ADG42685 standard; protein; 99 AA.
XX
AC      ADG42685;
XX
DT      26-FEB-2004 (first entry)
DE      Human PAGE-2 protein #4.
XX
KW      Human; NOVX; GAGE-like protein; interferon; GPCR; mast cell protease;
KW      hepatocyte nuclear factor; cancer; immune disorder; hepatic disorder;
KW      trauma; HIV; Parkinson's disease; Alzheimer's disease.
OS      Homo sapiens,
XX
PN      US2003202971-A1.
XX
PD      30-OCT-2003.
XX
PF      06-MAY-2002; 2002US-00139854.
XX
PR      15-FEB-2000; 2000US-0182723P.
PR      15-FEB-2000; 2000US-0182724P.
PR      15-FEB-2000; 2000US-0182733P.
PR      22-FEB-2000; 2000US-0183896P.
PR      23-FEB-2000; 2000US-0184275P.
PR      23-FEB-2000; 2000US-0184482P.
PR      24-FEB-2000; 2000US-0184497P.
PR      13-APR-2000; 2000US-0197083P.
PR      10-AUG-2000; 2000US-0224157P.
PR      18-SEP-2000; 2000US-0233405P.
PR      27-SEP-2000; 2000US-0236060P.
PR      02-JAN-2001; 2001US-0259414P.
PR      18-JAN-2001; 2001US-0262454P.
PR      14-FEB-2001; 2001US-00783429.
XX
PA      (MAUT/) MAJUMDER K.
XX
PI      Majumder K;
XX
DR      WPI; 2003-900615/82.
XX
PT      Novel isolated NOVX polypeptide useful for treating cancer, immune
PT      disorders, hepatic disorders, AIDS, Parkinson's disease.
XX
PS      Disclosure; SEQ ID NO 46; 155pp; English.
XX
XX      The invention relates to an isolated NOVX polypeptide comprising mature
XX      form of protein having a fully defined sequence ADG42655ADG42659, NOV8
XX      and NOV10, a variant of mature form of the protein, where any amino acid
XX      in mature protein changed to a different amino acid, provided that not
XX      more than 50% of amino acid in sequence of mature protein are so changed,
XX      or its fragment. Also included are an isolated nucleic acid molecule (NA)
XX      encoding NOVX (or portion or variant), a vector comprising NOVX NA, a
XX      cell comprising the vector, an anti-NOVX antibody, determining (M1) the
XX      presence or amount of NOVX NA in a sample, a pharmaceutical composition
XX      comprising NOVX, NOVX NA or the antibody and carrier) and a kit
XX      comprising in one or more containers. NOV8 and NOV10 are 2 of 13
XX      disclosed NOVX proteins showing homology to GAGE (G-antigen) proteins
XX      (NOV1-4), interferons (NOV5 and 6), G protein-coupled receptors (NOV7),
XX      Mast cell proteases (NOV9 and NOV10-13) and hepatocyte nuclear factors
XX      (NOV9). The antibody is useful for determining the presence or amount of
XX      NOVX in a sample. NOVX is useful for identifying an agent that binds to

```

```

CC      NOVX and for identifying a potential therapeutic agent for use in
CC      treatment of a pathology, where the pathology use related to aberrant
CC      expression or aberrant physiological interaction of NOVX. NOVX, NOVX NA
CC      or the antibody is useful for treating or preventing a pathology
CC      associated with NOVX e.g. cancer, immune disorder, hepatic disorder,
CC      trauma, HIV, Parkinson's disease, and Alzheimer's disease. The present
CC      sequence represents a protein showing sequence similarity to a NOVX
CC      protein (or fragment).
XX
SQ      Sequence 99 AA:
XX
Query Match      97.6%; Score 528; DB 7; Length 99;
Best Local Similarity 100.0%; Pred. No. 4e-48;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      4 RYRSRSRGRGDGEAPDVFAVAPGESQOEPPDNDIPRQEREGTPPIERKVEGDC 63
        1 RYRSRSRGRGDGEAPDVFAVAPGESQOEPPDNDIPRQEREGTPPIERKVEGDC 60
Db
QY      64 QEMDLEKTRSERGSDVYKKTPTPNPKAKTKRAGDGQP 102
        61 QEMDLEKTRSERGSDVYKKTPTPNPKAKTKRAGDGQP 99
Db
RESULT 7
ADJ55754
ID      ADJ55754 standard; peptide; 99 AA.
XX
AC      ADJ55754;
XX
DT      06-MAY-2004 (first entry)
DE      Peptide homologous to a fragment of a human NOV3 protein SegID 46.
XX
KW      human; NOVX; neurogenesis; cell differentiation; cell motility;
KW      cell proliferation; haematopoiesis; angiogenesis; proliferative disorder;
KW      cancer; mastocytosis; haematopoiesis; angiogenesis; proliferative disorder;
KW      viral infection; AIDS; hepatitis; neuro-olfactory; trauma; surgery;
KW      neoplastic; adenocarcinoma; lymphoma; asthma; Crohn's disease;
KW      multiple sclerosis; Alzright hereditary osteodystrophy; cystostatic;
KW      immunosuppressive; hepatotropic; virucide; anti-HIV; antiinflammatory;
KW      vulnery; antiaesthetic; neuroprotective.
XX
OS      Homo sapiens.
XX
PN      US2003199465-A1.
XX
PD      23-OCT-2003.
XX
PF      22-APR-2002; 2002US-00131409.
XX
PR      15-FEB-2000; 2000US-0182723P.
PR      15-FEB-2000; 2000US-0182724P.
PR      15-FEB-2000; 2000US-0182733P.
PR      22-FEB-2000; 2000US-0183896P.
PR      23-FEB-2000; 2000US-0184275P.
PR      23-FEB-2000; 2000US-0184482P.
PR      23-FEB-2000; 2000US-0184497P.
PR      24-FEB-2000; 2000US-0184744P.
PR      13-APR-2000; 2000US-0197083P.
PR      10-AUG-2000; 2000US-0224157P.
PR      18-SEP-2000; 2000US-0233405P.
PR      27-SEP-2000; 2000US-0236060P.
PR      02-JAN-2001; 2001US-0259414P.
PR      18-JAN-2001; 2001US-0262454P.
PR      14-FEB-2001; 2001US-00783429.
PR      03-JUL-2001; 2001US-00898953.
XX
PA      (MALY/) MALYANKAR U M.
PA      (TCHE/) TCHERNEV V T.
PA      (PADI/) PADIGARU M.
PA      (TAUP/) TAUPIER R J.

```

PA (SPYT/) SPYTEK K A.
 PA (GUX/) GUO X.
 PA (SPAD/) SPADERNA S K.
 PA (BOLD/) BOLDOG F L.
 PA (GERL/) GERLACH V.
 PA (ELLER/) ELLERMAN K.
 PA (MACD/) MACDOUGALL J R.
 PA (SMIT/) SMITHSON G.
 XX
 XX Maljankar UM, Tchernenov VT, Padigaru M, Taupier RJ, Spytek KA,
 PI Guo X, Spaderma SK, Boldog FL, Gerlach V, Ellerman K, Macdougall JR,
 PI Smithson G,
 XX
 DR WPI; 2003-900202/82.
 PT
 PT New polypeptides useful for treating e.g. cancer and mastocytosis, immune
 disorder, hepatic disorders, viral infections, asthma.
 XX
 PS Disclosure; SEQ ID NO 46; 86pp; English.
 XX
 CC This invention relates to a novel nucleic acids, and encoded polypeptides
 CC thereof, which have properties related to the stimulation of biochemical
 CC or physiological responses in a cell, tissue, organ or organism.
 CC Specifically, it refers to the use of biologically active fragments for
 CC diagnostic and prognostic assays and furthermore in the treatment of
 CC diverse pathological conditions. The present invention describes novel
 CC human NOVX proteins, as well as methods to identify small molecules that
 CC can be used to modulate or inhibit various processes including
 CC neurogenesis, cell differentiation, cell motility, cell proliferation,
 CC hemopoiesis and angiogenesis. Furthermore, they can be used in the
 CC manufacture of medicaments to treat proliferative disorders (e.g. cancer
 CC and mastocytosis), immune disorders, hepatic disorders (e.g. cancer
 CC viral infections) (e.g. AIDS and hepatitis), disorders of the neuro-
 CC olfactory system (e.g. trauma, surgery and/or neoplastic disorders),
 CC adenocarcinoma, lymphoma, asthma, Crohn's disease, multiple sclerosis and
 CC also for treating Alzhrigh hereditary osteodystrophy. Accordingly, they
 CC exhibit cytostatic, immunosuppressive, hepatotropic, virucide, anti-HIV,
 CC antiinflammatory, vulnerary, antiaesthetic and neuroprotective
 CC activities. This peptide is homologous to a fragment of a human NOVX
 CC protein of the invention.
 CC
 XX
 SO Sequence 99 AA;
 Query Match 97.6%; Score 528; DB 7; Length 99;
 Best Local Similarity 100.0%; Pred. No. 4e-48;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 RVSRSRGRGDDGQADPVVAFVAPGSGQSEPPPTNDIPEQERGTPIERKKGDC 63
 DB 1 RVSRSRGRGDDGQADPVVAFVAPGSGQSEPPPTNDIPEQERGTPIERKKGDC 60
 QY 64 QEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102
 DB 61 QEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 99
 RESULT 8
 ADM76592
 ID ADM76592 standard; protein; 99 AA.
 XX
 AC ADM76592;
 XX
 DT 03-JUN-2004 (first entry)
 DT
 DE Human NOV3 protein-related protein fragment SeqID46.
 XX
 KW NOVX; cytostatic; immunosuppressive; virucide; anti-HIV; hepatotropic;
 KW antiinflammatory; neuroprotective; nootropic; antiparkinsonian;
 KW gene therapy; vaccine; NOVX-associated disorder; cancer; immune disorder;
 KW viral infection; AIDS; hepatitis; neurological disorder;
 KW Alzheimer's disease; Parkinson's disease; olfactory disorder;
 KW chromosome mapping; tissue typing; NOV3; human.
 XX

OS Homo sapiens.
 XX
 XX WO2004009634-A2.
 XX
 XX 29-JAN-2004.
 PD
 XX
 PF 04-OCT-2001; 2001MO-US031284.
 XX
 XX 04-OCT-2000; 2000US-0237862P.
 PR 02-JAN-2001; 2001US-0259414P.
 PR 18-JAN-2001; 2001US-0262454P.
 PR 14-FEB-2001; 2001US-00783429.
 PR 03-JUL-2001; 2001US-00898953.
 PR 03-OCT-2001; 2001US-00970607.
 XX
 XX (CURA-) CURAGEN CORP.
 PA
 XX
 PI Gerlach VL, Ellerman K, Macdougall JR, Smithson G, Shinkets RA,
 PI Rastelli L, Herrmann J;
 XX
 DR WPI; 2004-143267/14.
 XX
 PT New NOVX polypeptides and nucleic acid molecules useful for diagnosing,
 PT preventing or treating NOVX-associated disorders (e.g. cancer, AIDS or
 PT Alzheimer's disease), and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 XX Disclosure; SEQ ID NO 46; 173pp; English.
 XX
 CC This invention relates to novel NOVX polypeptides and the DNA sequences
 CC which encode them. The invention may be useful for the development of
 CC compounds with a cytostatic, immunosuppressive, virucide, anti-HIV,
 CC hepatotropic, antiinflammatory, neuroprotective, nootropic or
 CC antiparkinsonian activity. In addition, the disclosed sequences may be
 CC useful for gene therapy or development of a vaccine. The NOVX
 CC polypeptide, nucleic acid or antibody is useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease. These
 CC are used in diagnosing, treating or preventing NOVX-associated disorders
 CC such as cancer, immune disorders, viral infections (for example AIDS or
 CC hepatitis), neurological disorders (for example Alzheimer's disease or
 CC Parkinson's disease) or olfactory disorders. The nucleic acids are
 CC further used as hybridisation probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. The present sequence
 CC is that of a human protein fragment which was used to demonstrate
 CC homology between the NOVX proteins of the invention and other human
 CC proteins.
 CC
 XX
 SO Sequence 99 AA;
 Query Match 97.6%; Score 528; DB 8; Length 99;
 Best Local Similarity 100.0%; Pred. No. 4e-48;
 Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 RVSRSRGRGDDGQADPVVAFVAPGSGQSEPPPTNDIPEQERGTPIERKKGDC 63
 DB 1 RVSRSRGRGDDGQADPVVAFVAPGSGQSEPPPTNDIPEQERGTPIERKKGDC 60
 QY 64 QEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 102
 DB 61 QEMDLEKTRSERGSDGVKKEKTPPNPKAKTKKAGDGP 99
 RESULT 9
 AAY12447
 ID AAY12447 standard; protein; 89 AA.
 XX
 AC AAY12447;
 XX
 DT 17-JUN-1999 (first entry)
 DT
 DE Human 5' EST secreted protein SEQ ID NO:478.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 XX

KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX
PN MO9906548-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB001222.
XX
PR 01-AUG-1997; 97US-00905135.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclet A, Lacroix B;
XX
DR WPI; 1999-153778/13.
XX
DR N-PSDB; AAX41280.
XX
PT New nucleic acids encoding human secreted proteins - obtained from cDNA
XX libraries prepared from e.g. liver, ovary, brain, prostate, kidney, lung,
XX umbilical cord, placenta and colon tissue.
XX
PS Claim 27, Page 778; 824pp; English.
XX
CC AAX41094 to AAX41347 represent 5' expressed sequence tags (ESTs) for
XX human secreted proteins, and encode the proteins given in AAY12261 to
XX AAY12514, respectively. The proteins given represent the signal peptide
XX and an N-terminal fragment of a secreted protein. The nucleic acid
XX sequences can be used for producing secreted human gene products. They
XX can also be used to develop products for diagnosis and therapy. The
XX proteins obtained may have cytokine activity, cell
XX proliferation/differentiation activity, haematopoiesis regulating
XX activity, tissue growth regulating activity, reproductive hormone
XX regulating activity, chemotactic/chemokinetic activity, haemostatic and
XX thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, tumour inhibition activity or other activities. The products
XX can be used in forensic, gene therapy and chromosome mapping procedures.
XX The sequences can also be used for obtaining corresponding promoter
XX sequences. The nucleic acids encoding the signal peptide can be used for
XX directing extracellular secretion of a polypeptide or the insertion of a
XX polypeptide into a membrane, or importing a polypeptide into a cell
XX
SQ Sequence 89 AA;
XX
Query Match 77.4%; Score 419; DB 2; Length 89;
Best Local Similarity 98.8%; Pred. No. 1.3e-36;
Matches 79; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 21 VVAFAVAGSGQEEPTNDIEPQERGTPIERKVEGDCQEMDLKTSERDGS 80
DB 8 VVAFVAPGSGQEEPTNDIEPQERGTPIERKVEGDCQEMDLKTSERDGS 67
QY 81 VKEKTPPNKAKTKRAGDG 100
DB 68 VKEKTPPNKAKTKRAGDG 87
XX
RESULT 10
AAY83167
ID AAY83167 standard; protein; 87 AA.
XX
AC AAY83167;
XX
DT 24-JUL-2000 (first entry)
XX
DE PAGE1 polypeptide.
XX
XX PAGE-4; MAGB; GAGE; reproduction; testis; prostate; fallopian tube;
KW

KW uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;
KW cytotoxic T lymphocyte; immune response; antibody; drug delivery;
KW immunconjugate.
XX
OS Homo sapiens.
XX
PN WO200012706-A1.
XX
PD 09-MAR-2000.
XX
PF 31-AUG-1999; 99WO-US020046.
XX
PR 01-SEP-1998; 98US-0098933P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Pastan I, Brinkmann U, Vasmatazis G, Lee B;
XX
DR WPI; 2000-237869/20.
XX
PT Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T
XX lymphocyte response and for raising antibodies which can be used to
XX detect the presence of PAGE-4 in cell samples or body tissues.
XX
PS Disclosure; Fig 1b; 63pp; English.
XX
CC PAGE-4 is a gene preferentially expressed in normal male and female
XX reproductive tissues e.g. prostate, testis, fallopian tube, uterus and
XX placenta, as well as in prostate cancer, testicular cancer and uterine
XX cancer. This expression pattern makes it a target for diagnosis and for
XX vaccine based therapy of such neoplasms. An isolated PAGE-4 peptide which
XX induces a cytotoxic T lymphocyte response when bound to a major
XX histocompatibility complex (MHC) class I molecule or the isolated PAGE-4
XX protein can be used in immunogenic compositions to raise a cytotoxic T
XX lymphocyte response against cells expressing PAGE-4 including cancer
XX cells of the prostate, uterus and testis. The nucleic acids encoding PAGE
XX -4 or PAGE-4 peptide fragments can also be used in these compositions.
XX Antibodies against PAGE-4 and its peptide fragments can be used in
XX detecting the presence of PAGE-4 and PAGE-4 coding sequences in cell
XX samples or body tissues. The presence of PAGE-4 in tissues which are not
XX related to reproduction can be indicative of the spread of cancerous
XX reproductive tissue. PAGE-4 can also be used to raise antibodies which
XX are then used as the targeting group of immunconjugates comprising
XX toxins used in therapeutic applications. This has applications for drug
XX delivery systems. This sequence is of the PAGE1 polypeptide which shares
XX sequence similarity with PAGE-4
XX
SQ Sequence 87 AA;
XX
Query Match 72.4%; Score 391.5; DB 3; Length 87;
Best Local Similarity 76.9%; Pred. No. 1e-33;
Matches 80; Conservative 3; Mismatches 2; Indels 19; Gaps 2;
QY 1 MSARVRSRSGRGQAEADVAVFAPGSGQEEPTNDQ--DIEPQERGTPIERK 58
DB 1 MSARVRSRSGRGQAEADVAVFAPGSGQEEPTNDQPDMEAFQ----- 48
QY 59 VEGDCQEMDLKTSERDGSVDYKETPPNPKAKTKRAGGCP 102
DB 49 -----QELDLKTSERDGSVDYKETPPNPKAKTKRAGGCP 87
XX
RESULT 11
ADJ33678
ID ADJ33678 standard; protein; 118 AA.
XX
AC ADJ33678;
XX
DT 22-APR-2004 (first entry)
XX
DE Protein of the invention SEQ ID NO:655.
XX
XX cycostatic; immunomodulator; antimicrobial; anti-inflammatory; vulnary;
KW

KM neuroprotective; gene therapy; cancer; bacterial infection;
 KM fungal infection; autoimmune disorder; inflammation; Parkinson's disease;
 KM allergy; wound healing; human.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 42 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

FT Misc-difference 103 /note= "Encoded by stop codon"

RESULT 12

AAV83168 standard; protein, 109 AA.

AAV83168;

24-JUL-2000 (first entry)

PAGE2 polypeptide.

PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;

uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;

cytotoxic T lymphocyte; immune response; antibody; drug delivery;

immunocjugate.

Homo sapiens.

MO200012706-A1.

09-MAR-2000.

31-AUG-1999; 99MO-US020046.

01-SEP-1998; 98US-0098993P.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Pastan I, Brinkmann U, Vasmatzis G, Lee B;

WPI; 2000-237869/20.

Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T

lymphocyte response and for raising antibodies which can be used to

detect the presence of PAGE-4 in cell samples or body tissues.

disclosure; Fig 1b; 63pp; English.

PAGE-4 is a gene preferentially expressed in normal male and female

reproductive tissues e.g. prostate, testis, fallopian tube, uterus and

placenta, as well as in prostate cancer, testicular cancer and uterine

cancer. This expression pattern makes it a target for diagnosis and for

vaccine based therapy of such neoplasms. An isolated PAGE-4 peptide which

induces a cytotoxic T lymphocyte response when bound to a major

histocompatibility complex (MHC) class I molecule or the isolated PAGE-4

protein can be used in immunogenic compositions to raise a cytotoxic T

lymphocyte response against cells expressing PAGE-4 including cancer

cells of the prostate, uterus and testis. The nucleic acids encoding PAGE

-4 or PAGE-4 peptide fragments can also be used in these compositions.

Antibodies against PAGE-4 and its peptide fragments can be used in

detecting the presence of PAGE-4 and PAGE-4 coding sequences in cell

samples or body tissues. The presence of PAGE-4 in tissues which are not

related to reproduction can be indicative of the spread of cancerous

reproductive tissue. PAGE-4 can also be used to raise antibodies which

are then used as the targeting group of immunocjugates comprising

toxins used in therapeutic applications. This has applications for drug

delivery systems. This sequence is of the PAGE2 polypeptide which shares

sequence similarity with PAGE-4

XX

Sequence 109 AA;

Query Match 54.6%; Score 295.5; DB 3; Length 109;

Best Local Similarity 61.5%; Pred. No. 26-23;

Matches 64; Conservative 4; Mismatches 31; Indels 5; Gaps 1;

10 RGRGDCQEA-PDVAVAFAPGE-----SQEPEPTNDIPGQREGTPTPEERKVEGDQCE 65

21 RGRGDCQEA-PDVAVAFAPGE-----SQEPEPTNDIPGQREGTPTPEERKVEGDQCE 80

66 MDLEKTRSERGSDGSDVEKTPPN-PKAKTKKAGDQCP 102

81 MDLEKTRSERGSDGSDVEKTPPN-PKAKTKKAGDQCP 118

RESULT 12

AAV83168 standard; protein, 109 AA.

AAV83168;

24-JUL-2000 (first entry)

PAGE2 polypeptide.

PAGE-4; MAGE; GAGE; reproduction; testis; prostate; fallopian tube;

uterus; placenta; cancer; major histocompatibility complex; MHC; CTL;

cytotoxic T lymphocyte; immune response; antibody; drug delivery;

immunocjugate.

Homo sapiens.

MO200012706-A1.

09-MAR-2000.

31-AUG-1999; 99MO-US020046.

01-SEP-1998; 98US-0098993P.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Pastan I, Brinkmann U, Vasmatzis G, Lee B;

WPI; 2000-237869/20.

Isolated PAGE-4 protein and peptide used for inducing a cytotoxic T

lymphocyte response and for raising antibodies which can be used to

detect the presence of PAGE-4 in cell samples or body tissues.

disclosure; Fig 1b; 63pp; English.

PAGE-4 is a gene preferentially expressed in normal male and female

reproductive tissues e.g. prostate, testis, fallopian tube, uterus and

placenta, as well as in prostate cancer, testicular cancer and uterine

cancer. This expression pattern makes it a target for diagnosis and for

vaccine based therapy of such neoplasms. An isolated PAGE-4 peptide which

induces a cytotoxic T lymphocyte response when bound to a major

histocompatibility complex (MHC) class I molecule or the isolated PAGE-4

protein can be used in immunogenic compositions to raise a cytotoxic T

lymphocyte response against cells expressing PAGE-4 including cancer

cells of the prostate, uterus and testis. The nucleic acids encoding PAGE

-4 or PAGE-4 peptide fragments can also be used in these compositions.

Antibodies against PAGE-4 and its peptide fragments can be used in

detecting the presence of PAGE-4 and PAGE-4 coding sequences in cell

samples or body tissues. The presence of PAGE-4 in tissues which are not

related to reproduction can be indicative of the spread of cancerous

reproductive tissue. PAGE-4 can also be used to raise antibodies which

are then used as the targeting group of immunocjugates comprising

toxins used in therapeutic applications. This has applications for drug

CC presence or amount of NOVX NA in a sample, a pharmaceutical composition
 CC (comprising NOVX, NOVX NA or the antibody and carrier) and a kit
 CC comprising in one or more containers. NOV8 and NOV10 are 2 of 13
 CC disclosed NOVX proteins showing homology to GAGE (G-antigen) proteins
 CC (NOV1-4), interferons (NOV5 and 6), G protein-coupled receptors (NOV7),
 CC Mast cell proteases (NOV8 and NOV10-13) and hepatocyte nuclear factors
 CC (NOV9). The antibody is useful for determining the presence or amount of
 CC NOVX in a sample. NOVX is useful for identifying an agent that binds to
 CC NOVX and for identifying a potential therapeutic agent for use in
 CC treatment of a pathology, where the pathology use related to aberrant
 CC expression or aberrant physiological interaction of NOVX. NOVX, NOVX NA
 CC or the antibody is useful for treating or preventing a pathology
 CC associated with NOVX e.g. cancer, immune disorder, hepatic disorder,
 CC trauma, HIV, Parkinson's disease, and Alzheimer's disease. The present
 CC sequence represents a protein showing sequence similarity to a NOVX
 CC protein (or fragment).
 XX
 SQ Sequence 109 AA;
 Query Match 54.6%; Score 295.5; DB 7; Length 109;
 Best Local Similarity 61.5%; Pred. No. 2e-23;
 Matches 64; Conservative 4; Mismatches 31; Indels 5; Gaps 1;
 QY 4 RYRSRSGRGDQGEADPVVAFVAPGE-----SQEPPPTNDIDIEGQREBGTPIPERK 58
 DB 6 RARSQSSERGNDESSQPVGSIVQEPTEKRQOEPPPTNDIDIEGQREBGTPIPERK 65
 QY 59 VEGDCQEMDLEKTRSERGSGSDYKEKTPPNPKAKTKKAGDGP 102
 DB 66 VEGDCQEMALKIETDEPGDGPVREGIMPTFPLTKVLEAGDAP 109
 RESULT 15
 ADJ55845 standard; peptide; 109 AA.
 XX
 AC ADJ55845;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Peptide homologous to a fragment of a human NOV1 protein SeqID 28.
 XX
 KW human; NOVX; neurogenesis; cell differentiation; cell motility;
 KW cell proliferation; hematopoiesis; angiogenesis; proliferative disorder;
 KW cancer; mastocytosis; immune disorder; hepatic disorder; cirrhosis;
 KW viral infection; AIDS; hepatitis; neuro-olfactory; trauma; surgery;
 KW neoplastic; adenocarcinoma; lymphoma; asthma; Crohn's disease;
 KW multiple sclerosis; Alibright hereditary osteodystrophy; cytoskeletal;
 KW immunosuppressive; hepatocytic; virulence; anti-HIV; anti-inflammatory;
 KW vulnerability; antiasthmatic; neuroprotective.
 XX
 OS Homo sapiens.
 XX
 PN US2003199465-A1.
 XX
 PD 23-OCT-2003.
 XX
 PF 22-APR-2002; 2002US-00131409.
 XX
 PR 15-FEB-2000; 2000US-0182723P.
 PR 15-FEB-2000; 2000US-0182724P.
 PR 15-FEB-2000; 2000US-0182733P.
 PR 22-FEB-2000; 2000US-0183966P.
 PR 23-FEB-2000; 2000US-0184275P.
 PR 23-FEB-2000; 2000US-0184482P.
 PR 23-FEB-2000; 2000US-0184497P.
 PR 24-FEB-2000; 2000US-0184744P.
 PR 13-APR-2000; 2000US-0197083P.
 PR 03-JUL-2000; 2000US-0215855P.
 PR 10-AUG-2000; 2000US-0224157P.
 PR 18-SEP-2000; 2000US-0233405P.
 PR 27-SEP-2000; 2000US-0236060P.
 PR 02-JAN-2001; 2001US-0259414P.

PR 18-JAN-2001; 2001US-0262454P.
 PR 14-FEB-2001; 2001US-00783429.
 PR 03-JUL-2001; 2001US-00898953.
 XX
 PA (MALY/) MALYANKAR U M.
 PA (TCHER) TCHERNEV V T.
 PA (PADI/) PADIGARU M.
 PA (TAUP/) TAUPIER R J.
 PA (SPYT/) SPYTEK K A.
 PA (GUOX/) GUO X.
 PA (SPAD/) SPADERNA S K.
 PA (BOLD/) BOLDOG F L.
 PA (GERL/) GERLACH V.
 PA (ELLER/) ELLERMAN K.
 PA (MACD/) MACDOUGALL J R.
 PA (SMIT/) SMITHSON G.
 XX
 PI Malynkar UM, Tchernev VT, Padigaru M, Spytex KA,
 PI Guo X, Spaderna SK, Boldog FL, Gerlach V, Ellerman K, Macdougall JR,
 PI Smithson G;
 DR WPI; 2003-900202/82.
 XX
 PT New polypeptides useful for treating e.g. cancer and mastocytosis, immune
 PT disorders, hepatic disorders, viral infections, asthma.
 XX
 PS Disclosure; SEQ ID NO 28; 86pp; English.
 XX
 CC This invention relates to a novel nucleic acids, and encoded polypeptides
 CC thereof, which have properties related to the stimulation of biochemical
 CC or physiological responses in a cell, tissue, organ or organism.
 CC Specifically, it refers to the use of biologically active fragments for
 CC diagnostic and prognostic assays and furthermore in the treatment of
 CC diverse pathological conditions. The present invention describes novel
 CC human NOVX proteins, as well as methods to identify small molecules that
 CC can be used to modulate or inhibit various processes including
 CC neurogenesis, cell differentiation, cell motility, cell proliferation,
 CC hematopoiesis and angiogenesis. Furthermore, they can be used in the
 CC manufacture of medicaments to treat proliferative disorders (e.g. cancer
 CC and mastocytosis), immune disorders, hepatic disorders (e.g. cirrhosis),
 CC viral infections (e.g. AIDS and hepatitis), disorders of the neuro-
 CC olfactory system (e.g. trauma, surgery and/or neoplastic disorders),
 CC adenocarcinoma, lymphoma, asthma, Crohn's disease, multiple sclerosis and
 CC also for treating Alibright hereditary osteodystrophy. Accordingly, they
 CC exhibit cytostatic, immunosuppressive, hepatocytic, virulence, anti-HIV,
 CC anti-inflammatory, vulnerability, antiasthmatic and neuroprotective
 CC activities. This peptide is homologous to a fragment of a human NOVX
 CC protein of the invention.
 XX
 SQ Sequence 109 AA;
 Query Match 54.6%; Score 295.5; DB 7; Length 109;
 Best Local Similarity 61.5%; Pred. No. 2e-23;
 Matches 64; Conservative 4; Mismatches 31; Indels 5; Gaps 1;
 QY 4 RYRSRSGRGDQGEADPVVAFVAPGE-----SQEPPPTNDIDIEGQREBGTPIPERK 58
 DB 6 RARSQSSERGNDESSQPVGSIVQEPTEKRQOEPPPTNDIDIEGQREBGTPIPERK 65
 QY 59 VEGDCQEMDLEKTRSERGSGSDYKEKTPPNPKAKTKKAGDGP 102
 DB 66 VEGDCQEMALKIETDEPGDGPVREGIMPTFPLTKVLEAGDAP 109

Search completed: August 24, 2005, 18:42:12
 Job time : 74 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 24, 2005, 18:35:22 ; Search time 21 Seconds

(without alignments)
467.339 Million cell updates/sec

Title: US-09-763-393-1

Perfect score: 541

Sequence: 1 MSARVRSRSGRGDQGEAPD.....EKTPPKAKTKKAGDGP 102

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	104.5	19.3	792	2	H83736	subtilisin-type a1
2	96.5	17.8	229	1	W4ML8	B4 protein - human
3	96.5	17.8	325	2	T18283	hypothetical prote
4	94.5	17.5	524	2	S35551	transcription fact
5	92.5	17.1	1791	2	T02345	hypothetical prote
6	92	17.0	606	2	S70358	centromere protein
7	89	16.5	396	1	A58938	surface protein rh
8	88	16.3	755	2	S32103	filensin - bovine
9	88	16.3	913	2	T52485	neurofilament prot
10	88	16.3	2109	2	E89066	protein H05009.1 (
11	88	16.3	2109	2	T33247	hypothetical prote
12	87.5	16.2	517	2	S20248	transcription fact
13	87	16.1	798	2	I50479	neurofilament medi
14	86	15.9	692	2	T13161	A-kinase anchor pr
15	85	15.7	649	2	D85135	hypothetical prote
16	85	15.7	896	2	T24169	hypothetical prote
17	85	15.7	906	2	T24166	hypothetical prote
18	85	15.7	5037	2	B35041	ryanodine receptor
19	84.5	15.6	359	2	S40746	hypothetical prote
20	84.5	15.6	1166	2	H86341	hypothetical prote
21	83.5	15.4	355	2	A33821	chromosomal protei
22	83.5	15.4	548	1	OPFCL	neurofilament tripl
23	83	15.3	837	2	T49562	neurofilament-H re
24	82.5	15.2	599	2	S18735	centromere protein
25	82.5	15.2	1271	2	T24008	hypothetical prote
26	82.5	15.2	5032	1	A35041	ryanodine receptor
27	82	15.2	541	2	T46423	hypothetical prote
28	82	15.2	610	2	T17775	protein kinase hom
29	82	15.2	690	2	H84309	protoporphyrin IX

30	81.5	15.1	247	2	A29545	submandibular glan
31	81.5	15.1	811	2	T04638	hypothetical prote
32	81.5	15.1	824	2	UC7532	cellulase (EC 3.2.
33	81.5	15.1	1192	2	A71623	probable secreted
34	81	15.0	316	2	JC6146	Card protein - Myx
35	81	15.0	952	2	S64473	translation initia
36	80.5	14.9	378	1	OZQOAB	circumsporozoite p
37	80.5	14.9	385	2	A48571	circumsporozoite p
38	80.5	14.9	646	1	S15901	chromogranin B pre
39	80.5	14.9	761	2	T32183	hypothetical prote
40	80.5	14.9	972	2	T49773	related to actin-1
41	80.5	14.9	2359	2	T03094	A-kinase anchor pr
42	80.5	14.9	5035	1	I46646	ryanodine receptor
43	80.5	14.9	5138	2	B96695	hypothetical prote
44	80	14.8	554	2	UW0094	neurofilament prote
45	80	14.8	2172	2	T20145	hypothetical prote

ALIGNMENTS

```

RESULT 1
H83736
subtilisin-type alkaline proteinase (EC 3.4.21.-) apr. precursor [similarity] - Bacillus
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: H83736
R/Takami, H.; Nakaoe, K.; Takaki, Y.; Maeno, G.; Sawaki, R.; Fujii, F.; Hara
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: A83650, MUID:20512582, PMID:11058132
A/Accession: H83736
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-792 <STO>
A/Cross-references: UNIPROT:Q9KFO3; GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA804
A/Experimental source: strain C-125
C/Genetics:
A/Gene: apr
C/Keywords: hydrolase; serine proteinase
F.1-28/Domain: signal sequence #status predicted <Sig>

Query Match      19.3%; Score 104.5; DB 2; Length 792;
Best Local Similarity 26.4%; Pred.No. 0.21;
Matches 29; Conservative 16; Mismatches 50; Indels 15; Gaps 3;

QY      7 SRSRGDQGEAPDVAFVAPGSGOQEPPTNODIEP--GOERGG-----TPPIERK 58
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      361 SSGHRLVGYGLAQYVSTPPDEENESPAENPQEQSDGKENSQSGSTPPDEEN 420
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      59 VEGDQENDLEKTRSERBGSDVKEKTPP-----NPKAKTKKAGDQ 101
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      421 EESPAEDQEQSDGKENSQSGSTPPDEENESPAENVQEQPSQK 470
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
W4ML8
B4 protein - human papillomavirus type 8
C/Species: human papillomavirus type 8
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C/Accession: A03677
R/Fuchs, P.G.; Ifthner, T.; Weninger, J.; Pfeister, H.
J. Virol. 58, 626-634, 1986
A/Title: Epidermodysplasia verruciformis-associated human papillomavirus 8: genomic seq
A/Reference number: A03019; MUID:86500410; PMID:3009874
A/Accession: A03677
A/Molecule type: DNA
A/Residues: 1-229 <FUC>
A/Cross-references: UNIPROT:P06425; GB:M12737; NID:G333074
A/Note: This ORF is not annotated in GenBank entry PRN8G
C/Superfamily: papillomavirus type 5 B4 protein
C/Keywords: early protein

```

```

Query Match          17.8%; Score 96.5; DB 1; Length 229;
Best Local Similarity 27.1%; Pred. No. 0.29;
Matches 29; Conservative 11; Mismatches 44; Indels 23; Gaps 3;

QY 6 RRSRGRGQGEADPVVAFVAPGESQOEPP-----TDNODIEPQEREG---TPPI 54
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 71 RKSRSDQGRDPAFGLAPGRSPGLAPPPYPGPPRRSPQFGPGPDDEGLQPL 130
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 55 EERKVGDDQEMDLEKTRSERGDGSVYKKEPPNPKAKTKKAGDQC 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 131 GGGQVEG-----HGGDQDPQGHPPPTPSNGKHGEBGDGE 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
T18283
hypothetical protein G5 - gline mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18283
R:Rieben Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyosawa, H.; Hugh
Genetics 148, 1117-1125, 1998
A:Title: Dictyostelium discoideum nuclear plasmid Ddb5 is a chimera related to the Ddb1
A:Reference number: Z14684; MUID:98198836; PMID:9539429
A:Accession: T18283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <R1E>
A:Cross-references: UNIPROT:O44016; EMBL:U00796; NID:G2702254; PID:G2702258; PIDN:AAC186
C:Genetics:
A:Introns: 85/1

Query Match          17.8%; Score 96.5; DB 2; Length 325;
Best Local Similarity 31.0%; Pred. No. 0.41;
Matches 27; Conservative 13; Mismatches 26; Indels 21; Gaps 3;

QY 29 ESQOEPP--TDNODIEPQEREGTPPIERKVGDCQEMDEKTR----- 72
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 EEEBETPAVSEBKEKEEBEETPAVSEBKEKEEBEEDKOKKEKIEBDTETGKKK 248
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 73 ----SRGSGDVYKKTTPNPKAKTK 95
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 DEWMERODGIRISEKTKDKPTE-KTK 274
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
S3551
transcription factor IIF chain RAP74 - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 09-Dec-1993 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S3551
R:Gong, D.W.; Hasegawa, S.; Wada, K.; Roeder, R.G.; Nakatani, Y.; Horikoshi, M.
Nucleic Acids Res, 20, 6736, 1992
A:Title: Cloning of three putative structural subdomains by comparison of primary st
A:Reference number: S3551; MUID:93126122; PMID:1480494
A:Accession: S3551
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-524 <GON>
A:Cross-references: UNIPROT:O04870; EMBL:Z17426; NID:G65029; PIDN:CAA78999.1; PID:G65030
C:Genetics:
A:Gene: RAP74
C:Keywords: DNA binding; transcription regulation

Query Match          17.5%; Score 94.5; DB 2; Length 524;
Best Local Similarity 31.5%; Pred. No. 0.99;
Matches 28; Conservative 15; Mismatches 33; Indels 13; Gaps 4;

QY 11 GGGDQGEADPVVAFVAPGESQOEPPPTNODIEPQEREGTPPIERKVGSCQEMDLEK 70
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 270 GPFEGQ-----VDYSDSSSDEELP--GKIKPAKEEGPKGLDQSSSS--ESEEK 320
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 71 TRSERGDSGVYKKTTPNPKAKTKKAGD 99
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 321 AEEERGE----EKKAPRPQDNKKKKGD 345
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
T02345
hypothetical protein KIAA0324 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02345
R:Ricke, D.O.; Bruce, D.; Mundt, M.; Doggett, N.; Munk, C.; Saunders, E.; Robinson, D.;
re, J.; White, S.; Ueng, S.; Tacum, O.; Campbell, C.; Fawcett, J.; Deaven, L.
submitted to the EMBL Data Library, March 1998
A:Description: Sequencing of human chromosome 16p13.3.
A:Reference number: Z14654
A:Accession: T02345
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1791 <R1C>
A:Cross-references: UNIPROT:O60382; EMBL:AC004493; NID:G2996648; PIDN:AAC08453.1; PID:G2
C:Genetics:
A:Map position: 16
A:Introns: 1610/2; 1706/2
A:Note: KIAA0324

Query Match          17.1%; Score 92.5; DB 2; Length 1791;
Best Local Similarity 25.5%; Pred. No. 5;
Matches 26; Conservative 17; Mismatches 38; Indels 21; Gaps 3;

QY 6 RRSRGRGQGEADPVVAF--VAPGESQOEPPPTNODIEPQEREGT----- 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 503 RTPSRGRSCDSSPEPKALPQTPRPRSRSPSPPELNKCLTPQERSGSSSDVKTVAR 562
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 52 PRIERKVGDCQEMDLEKTRSERGDGSVYKKEPPNPKAKTK 86
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 563 TPLGQRSSGSSQGLDVKPSASQERSSSDSSPDKAKTRTP 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
S70358
centromere protein B - Chinese hamster
C:Species: Crictetus griseus (Chinese hamster)
C:Date: 21-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S70358
R:Bejarano, L.A.; Valdivia, M.M.
Biochim. Biophys. Acta 1307, 21-25, 1996
A:Title: Molecular cloning of an intronless gene for the hamster centromere antigen CENP
A:Reference number: S70358; MUID:96254058; PMID:8652663
A:Accession: S70358
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-606 <BRJ>
A:Cross-references: UNIPROT:P48988; EMBL:U20951; NID:G836955; PIDN:AAB06494.1; PID:G8369
C:Keywords: DNA binding

Query Match          17.0%; Score 92; DB 2; Length 606;
Best Local Similarity 28.4%; Pred. No. 1.9;
Matches 27; Conservative 16; Mismatches 42; Indels 10; Gaps 2;

QY 8 RSRGRGQGEADPVVAFVAPGESQOEPPPTNODIEPQEREGTPPIERKVE-----G 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 REAGFGGLNNTITTSFSGSEEEEE---EEEEEEEEEGEGEEEEEEEEEGEG 440
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 DQEMDLEKTRSERGDGSVYKKTTPNPKAKTK 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 441 EGEEVGEEVEEVEEGDSEDEEEEEEEEEESSSE 475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
A58938
surface protein rhoptry ROP1 precursor - Toxoplasma gondii
C:Species: Toxoplasma gondii
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A58938; A56444; S37697

```

R.Boothroyd, J.C.
submitted to Genbank, July 1995
A:Reference number: A58938
A:Accession: A58938
A:Status: preliminary
A:Molecule type: RNA
A:Readsues: 1-396 <BOO>
A:Cross-references: UNIPROT:Q04151; GB:M1274; NID:g897822; PTDN:AAA69859.1; PTD:g897823
A>Note: revision to sequence reported in A45644
R.Osorio, P.N.; Schwartzman, J.D.; Boothroyd, J.C.
Mol. Biochem. Parasitol. 50, 1-15, 1992
A>Title: A Toxoplasma gondii rhoptry protein associated with host cell penetration has u
A:Reference number: A45644; PMID:52178277; PMID:1542304
A:Accession: A45644
A:Molecule type: mRNA
A:Readsues: 'MACROLCSQVNLFFLPRLDYCTDPEPT', 1-352, 'PEQR', 358-364, 'R', 366, 'I', 393, 'SP', 396
A:Cross-references: EMBL:M1274; NID:g897822
A>Note: sequence extracted from NCBI backbone (NCBIN:85178, NCBIIP:85179)
C:Superfamily: surface protein rhoptry
C:Keywords: surface antigen
F1-21/Domain: signal sequence #status predicted <SIG>
F12-396/Product: surface protein rhoptry #status predicted <MAT>
P12-396/Product: surface protein rhoptry #status predicted <MAT>

	Query Match	16.5%	Score 89;	DB 1;	Length 396;
	Best Local Similarity	30.2%;	Pred. No. 2.2;		
	Matches	Conservative 32;	Mismatches 37;	Indels 22;	Gaps 5;
Oy	16 QEAPYVAVFVARGESQ-----QEPPETDNDIPGQERESTPTIEE---RKVEG----	61			
Dd	143 QELPEDGLGVAPPNQAQLPAPPVQCLPPTBEQLLEPSTEQELAPPVAGEQRQLVGEHGCP	202			
Oy	62 -----DCQMDLKRSTRERGDSNVKELTPPRPKAKTKKEADGCP	102			
Dd	203 QGPYDDQQLLEPLT-EEQGEGP--QEPYPPPPPTREBQEGQP	245			

```

RESULT 8
S32103
filensin - bovine
N:Alternate names: Intermediate filament protein
C:Species: Bos primigenius taurus (cattle)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 10-Sep-1997
C:Accession: S32103; A40690
R:Gonnart, F.; Merdes, A.; Quinlan, R.; Hess, J.; Fitzgerald, P.G.; Ouzounis, C.; Georga
submitted to the EMBL Data Library, March 1993
A:Description: Bovine filensin possesses primary and secondary structure similarity to i
A:Reference number: S32103
A:Accession: S32103
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-755 <GOU>
A:Cross-references: EMBL:X72388; NID:g287751; PID:g287752
R:Gonnart, F.; Merdes, A.; Quinlan, R.; Hess, J.; Fitzgerald, P.G.; Ouzounis, C.A.; Geor
J. Cell Biol. 121, 847-853, 1993
A:Title: Bovine filensin possesses primary and secondary structure similarity to Interme
A:Reference number: A40690, MUID:93260017, PMID:8491777
A:Accession: A40690
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-622,'RP', 625-755 <G02>
A:Experimental source: lens
A:Note: sequence extracted from NCBI backbone (NCBIN:132495, NCBI:P:132499)
A:Note: part of this sequence was confirmed by protein sequencing
A:Keywords: membrane-associated protein

```

```

Query Match Similarity 16.3% Score 88; DB 2; Length 755;
Best Local Similarity 34.0%; Pred. No. 5;
Matches 33; Conservative 6; Mismatches 34; Indels 24; Gaps 6;

Qy 12 RGDGGEADPVVAFVAPGSGQEEPTDN-----QDIEQGQEHGTPPIERKRVKGDGCOEM 66
      |||
Db 496 KGDGDSVPPDSDGVSEPP--QQPEPLBEGCGQPPCKEKGGLKREGGP-----EG----- 541
      |||

```

OY 67 DEKTRSERGDSGVKEK -TPPNPKAKTKEAGDGP 102
| | | | |
DB 542 ---KGPEPEGKDSVKEEGSGPPEGGKGDVKEEG-GRP 574

RESULT 9
T52485
neurofilament protein NF-M(2), middle molecular weight [imported] - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 24-Oct-2000 #sequence_rev1sion 24-Oct-2000 #text_change 09-Jul-2004
C/Accession: T52485
R/Gervasi, C.; Szaro, B.G.
Brain Res. Mol. Brain Res. 48, 229-242, 1997
A/TITLE: Sequence and expression pattern of two forms of the middle molecular weight ne
A/Reference number: Z26090
A/Accession: T52485
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-913 <GB>
A/Cross-references: UNIPROT:O13099, EMBL:U85970, PIDD:AA853390.1

Query Match	16.3%	Score 88	DB 2	Length 913
Best Local Similarity	23.8%	Pred. No. 6.1		
Matches	31	Conservative	22	Mismatches 45; Indels 32; Gaps 5
QY	3	ARVRSRSGRGDDGOEA-----PDVVAFVAPGESOOERPEPTNDNDI-----	42	
DB	495	AAVSSSVQMAAPGEAAEGEBEEBEERKEGBEAAEBEGEBEEBKKEBGBDEBKYDDEBEG	554	
QY	43	-EPGOEREGTPIIEERKVEGDCQEMDLEKTRSERGD-----GSDVKEKTPPNPKAK---	93	
DB	555	GGGGBEEDVKAABEERKEBEGE-EDAAVEETVAEBGEBEGKGDABAKBEGBEEBKREV	613	
QY	94	--TKEAGDQ 101		
DB	614	AADEGEGEGE 623		

```

RESULT 10
E89066
protein H05009.1 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: E89066
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; NUID:99069613; PMID:9851916
A/Note: see webfics genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: E89066
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-2109 <STO>
A/Cross-References: UNIPROT:Q81SF7, UNIPROT:Q81SFE, UNIPROT:Q81SF5, GB:chr_V, PIDN:AAIC19
C/Genetics:
A/Gene: H05009.1
A/Map position: 5

```

Query Match	16.3%	Score 88	DB 2	Length 2109
Best Local Similarity	29.4%	Pred. No. 14		
Matches 32	Conservative 8	Mismatches 47	Indels 22	Gaps 3

QY	16	QEAADVVA-----FVAPGSGQEEPTTNDILRQGF--REGTF-----P	53
DB	1712	QEVMPVAAPEPIQEEVYPKEAAPEEPQSDVPKGAFLPEPQGVPPKAAPESGFTQSDVP	1771
QY	54	IEERKVGDCQEMDLEKTRSERGGSDVYKKEPTPNPKAKTKEAGDQOP	102
DB	1772	KEEAPSEPTQSDVPKEAAPSEPTQENVPKKAAPSEPTKDVYKEAAPSP	1820

RESULT 11


```

Db      582 ITAVRAVD---GGGAPAE-----SSGEPADDEGPTDT--AEAGSDPQAEQLLEEGVPC 631
QY      61 GDCQEMDLEKTRSERGSDVKEKTPPNPKAKTKEA 97
Db      632 GTAHEKGVPKARSEAEAGNGAETMAEAESAQTQVA 668

```

RESULT 15

```

D85135
hypothetical protein AT4g12610 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85135
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: D85135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-649 <STO>
A:Cross-references: UNIPROT:Q9SU25; GB:NC_001268; NID:g7267963; PIDN:CAB78304.1; GSPDB:C
C:Genetics:
A:Gene: AT4g12610
A:Map position: 4

```

```

Query Match      15.7%; Score 85; DB 2; Length 649;
Best Local Similarity 28.6%; Pred. No. 7.8;
Matches 26; Conservative 17; Mismatches 42; Indels 6; Gaps 3;

```

```

QY      14 DGGEAPDVVA--FVAPESSQGEPPPTDNDIEPQGEREGTPPIERKRVESDCCQEMDLEKT 71
Db      303 DPBEREDLLAPEIPAPEPIKQDEBDENEEEGGLSKSGK---ELKKLLGKANGLD-ESD 358
QY      72 RSERGDSDVKEKTPPNPKAKTKEAGDGGP 102
Db      359 EDDDDSDDEETNYGTVNSKQENAKKEP 389

```

Search completed: August 24, 2005, 18:43:43
 Job time : 23 secs

This Page Blank (uspto)

Result	Query	Match	length	DB	ID	Description
No.	Score					
1	181.5	33.5	11.1	4	US-09-621-976-7591	Sequence 7591, App
2	170	31.4	11.0	4	US-09-621-976-7592	Sequence 7592, App
3	122	22.6	11.6	4	US-09-163-748C-7	Sequence 8, App11-8
4	107	19.8	11.7	4	US-09-163-748C-7	Sequence 7, App11-8
5	107	19.8	86.1	4	US-09-873-737A-6	Sequence 6, App11-8
6	106.5	19.7	86.2	4	US-09-873-737A-4	Sequence 4, App11-8
7	93.5	17.3	33.2	3	US-08-301-162-2	Sequence 2, App11-8
8	93.5	17.3	33.2	3	US-09-461-240-2	Sequence 2, App11-8
9	93.5	17.3	33.2	4	US-09-968-927-2	Sequence 2, App11-8
10	93.5	17.3	42.8	3	US-08-301-162-18	Sequence 18, App11-8
11	93.5	17.3	42.8	3	US-09-461-240-18	Sequence 18, App11-8
12	93.5	17.3	42.8	4	US-09-968-927-18	Sequence 18, App11-8
13	93.5	17.3	66.7	3	US-09-303-064-55	Sequence 55, App11-8
14	93.5	17.3	66.7	3	US-09-086-503-55	Sequence 55, App11-8
15	87.5	16.2	51.7	4	US-09-949-016-6170	Sequence 6170, App11-8
16	87.5	16.2	52.0	4	US-09-949-016-7728	Sequence 7728, App11-8
17	85	15.7	50.7	4	US-09-424-783-4	Sequence 4, App11-8
18	84.5	15.6	20.79	4	US-09-949-016-8301	Sequence 8301, App11-8
19	84	15.5	43.7	4	US-09-252-991A-28176	Sequence 28176, App11-8
20	82.5	15.2	53.9	4	US-09-538-092-864	Sequence 864, App11-8
21	82.5	15.2	50.32	4	US-09-538-092-979	Sequence 979, App11-8
22	82	15.2	70.1	4	US-09-252-991A-24048	Sequence 24048, App11-8
23	81.5	15.1	56.5	4	US-09-252-991A-20122	Sequence 20122, App11-8
24	81	15.0	26.3	4	US-09-902-540-11171	Sequence 11171, App11-8
25	80.5	14.9	23.0	4	US-09-248-796A-14401	Sequence 14401, App11-8
26	80.5	14.9	62.8	4	US-09-252-991A-24491	Sequence 24491, App11-8
27	80	14.8	13.9	4	US-09-270-767-36313	Sequence 36313, App11-8

28	80	14.8	159	4	US-09-02-270-767-51530	Sequence 51530, A
29	80	14.8	339	4	US-09-02-252-991A-23619	Sequence 23619, A
30	80	14.8	396	4	US-09-06-940-211A-1055	Sequence 1055, A
31	-80	14.8	519	4	US-09-09-602-540-15378	Sequence 15378, A
32	79.5	14.7	232	6	5171843-7	Patient No. 5171843
33	79.5	14.7	232	6	5171843-7	Patient No. 5171843
34	79.5	14.7	378	6	5171843-9	Patient No. 5171843
35	79.5	14.7	378	6	5171843-9	Patient No. 5171843
36	79	14.6	775	4	US-09-09-902-540-12637	Sequence 12637, A
37	79	14.6	843	4	US-09-08-873-737A-12	Sequence 2, Appt1
38	78.5	14.5	506	2	US-08-08-820-170A-19	Sequence 19, Appt1
39	78.5	14.5	506	3	US-09-05-659-19	Sequence 19, Appt1
40	78.5	14.5	506	3	US-09-02-273-565-19	Sequence 19, Appt1
41	78.5	14.5	506	3	US-09-06-553-538-19	Sequence 19, Appt1
42	78.5	14.5	506	3	US-09-06-661-468-19	Sequence 19, Appt1
43	78.5	14.5	506	4	US-09-09-976-155-19	Sequence 19, Appt1
44	78.5	14.5	506	4	US-09-09-934-967-5	Sequence 1381, Appt1
45	78.5	14.5	758	4	US-09-09-934-967-5	Sequence 5, Appt1

ALIGNMENTS

```

RESULT 1
US-09-621-976-7591
: Sequence 7591, Application US/09621976
: Patent No. 6635053
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTs and Encoded Human Proteins
: FILE REFERENCE: GENSET 054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 7591
: LENGTH: 111
: TYPE: prt
: ORGANISM: Homo sapiens
US-09-621-976-7591

```

Query Match	33.5%	Score 181.5	DB 4	Length 11
Best Local Similarity	44.1%	Pred. NO. 3e-12		
Matches 49		Conservative 11	Mismatches 42	Indels 9
				Gaps 4

QY MSARVRSRSG--RGCGEAPDVVAEAVGGE-----SOOEPTNDQDIIEP-CQ-EREGT 51
 1 MSHVATRRSSSRGNDQSSQPVGSVIYQEPLEEKRQGEPEPTNDQGLAFSEIINEGA 60

Db

QY PPIEEKRVEDDCEQMDLEKTRSEKRGSGDYVEXETPPNPKAKTKKAGDGP 102
 61 PAVQGPDMKPVQDELALTKIIDEKPGGPVDRBSIEMTFDLTKYLEKGDNP 111

```

RESULT 2
US-09-621-976-7592
? Sequence 7592, Application US/09621976
? Patent No. 6639063
? GENERAL INFORMATION:
? APPLICANT: Dumas Milne Edwards, J.B.
? APPLICANT: Jobert, S.
? APPLICANT: Giordano, J.Y.
? TITLE OF INVENTION: ESTs and Encoded Human Proteins
? FILE REFERENCE: GENSET.054PR2
? CURRENT APPLICATION NUMBER: US/09/621.976
? NUMBER OF SEQ ID NOS: 2000-07-21
? SOFTWARE: Patent.pgm
? SEQ ID NO 7592
? LENGTH: 110
? TYPE: PRT

```

```

01 49  EGTPPIEERKVEDCOEMDEKTRSRERGSGSVKERTPNPFGHATKXAGDQ 101
02 61  EGASAGGKPRFADSQEGGHPTGTCECECDGPDGQEMDPENPEEVXTPBEKXQ 113
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1
```

27 QFEDVEVEMIEPATPEEGEPATRODPAHAQSEDEBGSAGGQPHQTKPEAHSEGGC 86
69 EKTRSERGDSDVYKENTPPNPKAKTKKEDCQ 101
87 ECE----DGPDOGEMDPNPEVVKTPBEGEKO 114

RESULT 5
US-09-873-737A-6
US-Sequence 6, Application US/09873737A

```

1  APPLICANT: Duke University
2  APPLICANT: Lin, Haifan
3  TITLE OR INVENTION: PURIFIED AND ISOLATED p1v1 FAMILY GENES AND GENE
4  TITLE OR INVENTION: PRODUCTS AND THERAPEUTIC AND SCREENING METHODS USING SAME
5  FILE REFERENCE: Attorney Docket No. 6723534 180-104/2
6  CURRENT APPLICATION NUMBER: US/09/873,737A
7  CURRENT FILING DATE: 2001-06-04
8  PRIOR APPLICATION NUMBER: PCT/US99/28764
9  PRIOR FILING DATE: 1999-12-03
10 PRIOR APPLICATION NUMBER: 60/110,901
11 PRIOR FILING DATE: 1998-12-04
12 NUMBER OF SEQ ID NOS: 21
13 SOFTWARE: PatentIn Ver. 2.1
14 SEQ ID NO 6
15 LENGTH: 861
16 TYPE: PRT
17 ORGANISM: Homo sapiens
18 FEATURE:
19 NAME/KEY: misc_feature
20 LOCATION: (76)

```

```

; NAME/KEY: misc_feature
; LOCATION: (303)
; OTHER INFORMATION: Xaa=Leu or Ile
; NAME/KEY: misc_feature
; LOCATION: (725)
; OTHER INFORMATION: Xaa=Leu or Ile
; US-09-873-73A-6
```

Query Match 19.8%; Score 107; DB 4; Length 861;
Best Local Similarity 28.3%; Pred. No. 0.0039;
Matches 30; Conservative 18; Mismatches 50; Indels 8; Gaps 3;
QY 1 MSARVSRSRGRDGOEAPDVAFVA---PG--ESQOEPPPTNODIEPGQER---EGTP 52
1 MTGRARARARAGRGAGTQAGTAVSTAGQCGYIQPRPQPPAGGELFGRGRGRTAGTA 60
DB 53 PIERKVEGDCQEMDLEKTRSERGDSGVYEKTPPNPKAKTKEAG 98
61 KSGQLQISAGFQELISXNERGRRRDPHDLGVNTRONLDHVSKSTG 106
RESULT 6
US-09-873-737A-4
Sequence 4, Application US/09873737A
Patent No. 6723534
GENERAL INFORMATION:
APPLICANT: Duke University
APPLICANT: Ldn, Halfan
TITLE OF INVENTION: PURIFIED AND ISOLATED p1w1 FAMILY GENES AND GENE
FILE REFERENCE: ACTOINNEY DOCKET NO. 6723534 180-104/2
CURRENT APPLICATION NUMBER: US/09/873,737A
CURRENT FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: PCT/US99/28764
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: 60/110,901
PRIOR FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 862
TYPE: PRT
ORGANISM: Mus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (90)
OTHER INFORMATION: Xaa=Leu or Ile
NAME/KEY: misc_feature
LOCATION: (216)
OTHER INFORMATION: Xaa=Leu or Ile
NAME/KEY: misc_feature
LOCATION: (383)
OTHER INFORMATION: Xaa=Leu or Ile
NAME/KEY: misc_feature
LOCATION: (916)
OTHER INFORMATION: Xaa=Leu or Ile
US-09-873-737A-4
Query Match 19.7%; Score 106.5; DB 4; Length 862;
Best Local Similarity 29.1%; Pred. No. 0.0043;
Matches 32; Conservative 15; Mismatches 48; Indels 15; Gaps 3;
QY 1 MSARVSRSRGRDGOEAPDVAFVAPGSGQ-----EPPPTNODIEPGQER--- 48
1 MTGRARARARAGRGAGTQAGTAVSTAGQCGYIQPRPQPPAGGELFGRGRGRTAGTA 57
DB 49 EGTPPIERKVEGDCQEMDLEKTRSERGDSGVYEKTPPNPKAKTKEAG 98
58 GATSKSGQLQISAGFQELISXNERGRRRDPHDLGVNTRONLDHVSKSTG 107
RESULT 7
US-08-301-162-2
Sequence 2, Application US/08301162
Patent No. 6022546
GENERAL INFORMATION:
APPLICANT: Knapp, Stefan
APPLICANT: Ziegelmaler, Robert
APPLICANT: Kupper, Hans
TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
PREPARATION THEREOF AND THE USE THEREOF

NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/301,162
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/167,128
FILING DATE:
APPLICATION NUMBER: US 07/623,086
FILING DATE: 06-DEC-1990
APPLICATION NUMBER: DE P3940598.2
FILING DATE: 08-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: Fleisher, Raz E.
REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 02481.1005-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 392 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-301-162-2
Query Match 17.3%; Score 93.5; DB 3; Length 392;
Best Local Similarity 30.3%; Pred. No. 0.041;
Matches 30; Conservative 15; Mismatches 39; Indels 15; Gaps 4;
QY 16 QEAPDVAFVAPGSGQEEPPPTNODIEPGQEREGTPPIE---RKVEG-----DC 63
190 QELPPTGELPPTGELPPTGELPPTGELPPTGELPPTGELPPTGELPPTGELPPTG 249
DB 64 QEMDLEKTRSERGDSGVYEKTPPNPKAKTKEAGDGP 102
250 QQLLEPT-EEQOEGP--QELPPTGELPPTGELPPTGELPPTGELPPTGELPPTG 285
RESULT 8
US-09-461-240-2
Sequence 2, Application US/09461240
Patent No. 6326008
GENERAL INFORMATION:
APPLICANT: Knapp, Stefan
APPLICANT: Ziegelmaler, Robert
APPLICANT: Kupper, Hans
TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
PREPARATION THEREOF AND THE USE THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:

```

1      MEDIUM TYPE: Floppy disk
2      COMPUTER: IBM PC compatible
3      OPERATING SYSTEM: PC-DOS/MS-DOS
4      SOFTWARE: Patent In Release #1.0, Version #1.25
5
6      CURRENT APPLICATION DATA:
7      APPLICATION NUMBER: US/09/461,240
8      FILING DATE: 16-Dec-1999
9      CLASSIFICATION: <Unknown>
10
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: US/08/301,162
13     FILING DATE: <Unknown>
14     APPLICATION NUMBER: US/08/167,128
15     FILING DATE: <Unknown>
16     APPLICATION NUMBER: US 07/623,086
17     FILING DATE: 06-DEC-1990
18     APPLICATION NUMBER: DE P3940598.2
19     FILING DATE: 08-DEC-1989
20
21     ATTORNEY/AGENT INFORMATION:
22     NAME: Fleisher, Raz E.
23     REGISTRATION NUMBER: 34,131
24     REFERENCE/DOCKET NUMBER: 02481.1005-00000
25
26     TELECOMMUNICATION INFORMATION:
27     TELEPHONE: 202-408-4000
28     TELEFAX: 202-408-4400
29
30     INFORMATION FOR SEQ ID NO: 2:
31     SEQUENCE CHARACTERISTICS:
32     LENGTH: 392 amino acids
33     TYPE: amino acid
34     TOPOLOGY: linear
35
36     MOLECULE TYPE: peptide
37
38     SEQUENCE DESCRIPTION: SEQ ID NO: 2:
39
40     US-09-461-240-2

```

```

Query Match          17.3%; Score 93.5; DB 3; Length 392;
Beet Local Similarity 30.3%; Pred. No. 0.04; Indels 15; Gaps 4
Matches 30; Conservative 15; Mismatches 39; Indels 15; Gaps 4

QY      16 QEADVAVFVAFVAGSESQQEEPTNDODIPGQEREGTPIIE---RKVEG-----DC 63
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      190 QELPPPTQDELPPRTPEQELPPTTBELAPSTQEQLPPVGEGQRQLQVRHEHQPQGPPYDD 249
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      64 QEMDLEKTRSERGDGSDVKEKTPPNPKAKTKAAGDGOP 102
           :: | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      250 QQLLLEPT-EEQQEGEP--QEPLPPPPPTPRGGQPGPGQP 285
           :: | | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
US-09-968-927-2
Sequence 2, Application US/09968927
Patent No. 6419925

GENERAL INFORMATION:
APPLICANT: Knapp, Stefan
              Ziegelmaier, Robert
              Kupper, Hans
TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
                  Preparation Thereof and the Use Thereof
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
                Dunner
                STREET: 1100 I Street, N.W., Suite 700
                CITY: Washington
                STATE: D.C.
                COUNTRY: USA
                ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/968,927
FILING DATE: 03-Oct-2001

```

```

1 CLASSIFICATION: <Unknown>
2
3 PRIOR APPLICATION DATA:
4
5 APPLICATION NUMBER: US/08/301,162
6
7 FILING DATE: <Unknown>
8
9 APPLICATION NUMBER: US/08/167,128
10
11 FILING DATE: <Unknown>
12
13 APPLICATION NUMBER: US 07/623,086
14
15 FILING DATE: 06-DEC-1990
16
17 APPLICATION NUMBER: DE P3940598.2
18
19 FILING DATE: 08-DEC-1989
20
21 ATTORNEY/AGENT INFORMATION:
22
23 NAME: Flashner, Raz E.
24
25 REGISTRATION NUMBER: 34,331
26
27 REFERENCE/DOCKET NUMBER: 02481.1005-000000
28
29 TELECOMMUNICATION INFORMATION:
30
31 TELEPHONE: 202-408-4000
32
33 TELEFAX: 202-408-4400
34
35 INFORMATION FOR SEQ ID NO: 2:
36
37 SEQUENCE CHARACTERISTICS:
38
39 LENGTH: 392 amino acids
40
41 TYPE: amino acid
42
43 TOPOLOGY: linear
44
45 MOLECULE TYPE: peptide
46
47 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
48
49 US-09-968-927-2

```

```

Query Match 17.3% Score 93.5; DB 4; Length 392;
Best Local Similarity 30.3%; Pred. No. 0.041;
Matches 30; Conservative 15; Mismatches 39; Indels 15; Gaps 4

          Oy      Db      Oy      Db
          16      190      64      250
OQAPDVAVFAPGSGQEEPTNODIEPQGRGTPPIE---RKVEG-----DC 63
DBEPPTEGELPPTEGELPPTEGELAPSTEGELPPVSGORLQVPGHGPGQPPYD 249
OEMDLEKTRSGRGSDVKEKTPPPPKAKTKKAGDGP 102
|::|||:::|::|::|::|
QQLLEPT-EEQEGP--QEPUPPPPTTREGQEGGP 285

RESULT 10
US-08-301-162-18
: Sequence 18, Application US/08301162
: Patent No. 6022546
: GENERAL INFORMATION:
: APPLICANT: Knapp, Stefan
: APPLICANT: Ziegelmaier, Robert
: APPLICANT: Kupper, Hans
: TITLE OF INVENTION: Toxoplasma Gondii Antigens, The
: TITLE OF INVENTION: Preparation Thereof and the Use Thereof
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
: ADDRESSEE: Dunner
: STREET: 1300 I Street, N.W., Suite 700
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/301,162
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/167,128
: FILING DATE:
: APPLICATION NUMBER: US 07/623,086
: FILING DATE: 06-DEC-1990
: APPLICATION NUMBER: DE P3940598.2

```


US-09-968-927-18

Query Match 17.3%; Score 93.5; DB 4; Length 428;
Best Local Similarity 30.3%; Pred. No. 0.046;
Matches 30; Conservative 15; Mismatches 39; Indels 15; Gaps 4;

QY 16 QEAPDVVAFAVAGSQQEPPPTDNDIEFGQEREGTPIIE--RKVEG-----DC 63
DB 182 QELPPTEQELPPPTQELPPPTQELPSTQELPPVGEQORLQVPEHGPQGPYYD 241

QY 64 QEMDEKTRSERGDSGVKEKTPPNPKHAKTKKAGDGP 102
DB 242 QQLLEPT-EEQOEGP--QEPUPPPPPPTRGEGPBGQGP 277

RESULT 13
US-09-303-064-55
Sequence 55, Application US/09303064

Patent No. 6221619
GENERAL INFORMATION:
APPLICANT: MAINE, Gregory T.
APPLICANT: HUNT, Jeffery C.
APPLICANT: BROJANAC, Susan
APPLICANT: JYH-TSING SHEU, Michael
APPLICANT: CHOVAN, Linda E.
APPLICANT: TYNER, Joan D.
APPLICANT: HOWARD, Lawrence V.
APPLICANT: PARMELEY, Stephen F.
APPLICANT: REMINGTON, Jack S.
APPLICANT: ARABUO, Fausto
APPLICANT: SUZUKI, Yasuhiko
APPLICANT: LI, Shuli
TITLE OF INVENTION: ANTIGEN COCKTAILS, P35 AND USES THEREOF
FILE REFERENCE: 6361.US.P1
CURRENT APPLICATION NUMBER: US/09/303,064
CURRENT FILING DATE: 1999-04-30
EARLIER APPLICATION NUMBER: 09/086,503
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 55
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 55
LENGTH: 667
TYPE: PRT
ORGANISM: Toxoplasma gondii
US-09-303-064-55

Query Match 17.3%; Score 93.5; DB 3; Length 667;
Best Local Similarity 30.3%; Pred. No. 0.079;
Matches 30; Conservative 15; Mismatches 39; Indels 15; Gaps 4;

QY 16 QEAPDVVAFAVAGSQQEPPPTDNDIEFGQEREGTPIIE--RKVEG-----DC 63
DB 329 QELPPTEQELPPPTQELPPPTQELPSTQELPPVGEQORLQVPEHGPQGPYYD 388

QY 64 QEMDEKTRSERGDSGVKEKTPPNPKHAKTKKAGDGP 102
DB 389 QQLLEPT-EEQOEGP--QEPUPPPPPPTRGEGPBGQGP 424

RESULT 14
US-09-086-503-55
Sequence 55, Application US/09086503A
Patent No. 6129157
GENERAL INFORMATION:
APPLICANT: MAINE, Gregory T.
APPLICANT: HUNT, Jeffery C.
APPLICANT: BROJANAC, Susan
APPLICANT: JYH-TSING SHEU, Michael
APPLICANT: CHOVAN, Linda E.
APPLICANT: TYNER, Joan D.
APPLICANT: HOWARD, Lawrence V.
TITLE OF INVENTION: ANTIGEN COCKTAILS AND USES THEREOF
FILE REFERENCE: 6361.US.01

CURRENT APPLICATION NUMBER: US/09/086,503A
CURRENT FILING DATE: 1998-05-28

NUMBER OF SEQ ID NOS: 55
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 55
LENGTH: 667
TYPE: PRT
ORGANISM: Toxoplasma gondii
US-09-086-503-55

Query Match 17.3%; Score 93.5; DB 3; Length 667;
Best Local Similarity 30.3%; Pred. No. 0.079;
Matches 30; Conservative 15; Mismatches 39; Indels 15; Gaps 4;

QY 16 QEAPDVVAFAVAGSQQEPPPTDNDIEFGQEREGTPIIE--RKVEG-----DC 63
DB 329 QELPPTEQELPPPTQELPPPTQELPSTQELPPVGEQORLQVPEHGPQGPYYD 388

QY 64 QEMDEKTRSERGDSGVKEKTPPNPKHAKTKKAGDGP 102
DB 389 QQLLEPT-EEQOEGP--QEPUPPPPPPTRGEGPBGQGP 424

RESULT 15
US-09-949-016-6170
Sequence 6170, Application US/09949016

Patent No. 6812319
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6170
LENGTH: 517
TYPE: PRT
ORGANISM: Human
US-09-949-016-6170

Query Match 16.2%; Score 87.5; DB 4; Length 517;
Best Local Similarity 29.9%; Pred. No. 0.26;
Matches 35; Conservative 11; Mismatches 44; Indels 27; Gaps 6;

QY 9 SRGRDGEAPDVVAFAVAGSQQEPP--TDNDIEFGQEREGTPIIEKKVEGD-----DC 62
DB 275 SDGSSSSQEEBESKA-KAP--QOEGPKGVDEQSDSEEBEKKPPEDDEEBEKKAP 330

QY 63 -QEMDEKTRSERGDSG-----VEKTPP--NPCHAKTKKAGDGP 102
DB 331 TPQEKRRKDSSEESDSSEESDISEASSAFPMKKKTPPKRERKPSGSSRGNSRP 387

Search completed: August 24, 2005, 18:47:22
Job time : 44 secs